

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 16:09:21 : Search time 71.77 Seconds
(Without alignments)
253.023 Million cell updates/sec

Title: US-09-052-089a-1

Sequence: 1 MPRLCTICSDFDHSDV.....YRKTVPSLFOAKLDTFLMS 469

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match length	ID	Description
1	192.5	8.1	1938	1 MYHD_HUMAN	09UKX3 homo sapien
2	187.5	7.9	2663	1 CENE_HUMAN	002224 homo sapien
3	183.5	7.7	1957	1 YD86_SCHPO	Q10411 schizosacch
4	182.5	7.7	1290	1 XCPX_XENLA	P50332 xenopus lae
5	178	7.5	962	1 VDP_HUMAN	060763 homo sapien
6	176	7.4	1509	1 MYSN_ACACA	P05659 acanthameb
7	173.5	7.3	1935	1 MYSS_CTECA	090339 cyprinus ca
8	172	7.2	1938	1 MYH4_RABIT	Q28641 oryctolagus
9	172	7.2	2116	1 MYH2_DICDI	P08739 dictyosteli
10	170	7.1	978	1 RASO_AQUAE	067124 aquifex aeo
11	169.5	7.1	1744	1 TANA_XENLA	001550 xenopus lae
12	169.5	7.1	1940	1 MYH3_RAT	P12847 rattus norv
13	168	7.1	1935	1 MYH7_RAT	P02564 rattus norv
14	168	7.0	941	1 VDP_MOUSE	Q02150 mus musculi
15	168	7.0	1325	1 G160_MOUSE	P55937 mus musculi
16	168	7.0	1941	1 MYH2_HUMAN	09UKX2 homo sapien
17	167	7.0	879	1 MYSP_ONCVO	002171 onchocerca
18	167	7.0	880	1 MYSP_BRUMA	Q01202 brugia mala
19	167	7.0	2411	1 MYSA_DROME	P05661 drosophila
20	166	7.0	388	1 PAM_STRPY	P01542 streptococ
21	166	7.0	959	1 VDP_RAT	P15243 sus scrofa
22	166	7.0	1935	1 MYH7_PIG	P19293 sus scrofa
23	166	7.0	1939	1 MYH1_HUMAN	P35579 homo sapien
24	166	7.0	1960	1 MYH9_HUMAN	Q03001 homo sapien
25	166	7.0	2704	1 BPAL_HUMAN	P25366 saccharomyc
26	165.5	6.9	1790	1 USOI_YEAST	Q29230 archaeoglob
27	165	6.9	848	1 MYSP_DIRIM	099208 pyrococcus
28	165	6.9	886	1 RASO_ARCFU	P08089 streptococ
29	164.5	6.9	880	1 RASO_PVRAB	Q9L423 bos taurus
30	163.5	6.9	483	1 M6_STRPY	P15215 drosophila
31	163.5	6.9	1453	1 Y373_BOVIN	Q9UKT5 schizosacch
32	163.5	6.9	1639	1 LMGI_DROME	
33	163.5	6.9	1727	1 ALMI_SCHPO	

34	163.5	6.9	1976	1 MYH4_HUMAN	P35580 homo sapien
35	163.5	6.9	2871	1 DESP_HUMAN	P15924 homo sapien
36	163	6.8	1935	1 MYH7_HUMAN	P12883 homo sapien
37	163	6.8	1938	1 MYSS_CHICK	P13538 gallus gall
38	162.5	6.8	1232	1 KF4A_HUMAN	Q95239 homo sapien
39	162.5	6.8	1938	1 MYA_AEQIR	P24733 aequipecten
40	162.5	6.8	3911	1 AKAA_HUMAN	Q99996 h a-kinase
41	162	6.8	963	1 KINH_HUMAN	P33176 homo sapien
42	162	6.8	1966	1 MYSP_CAEEL	P02566 caenorhabdi
43	161.5	6.8	1939	1 MYH4_HUMAN	Q9623 homo sapien
44	161.5	6.8	1969	1 MYSA_CAEEL	P12844 caenorhabdi
45	161.5	6.8	1972	1 MYHB_MOUSE	Q08638 mus musculi

ALIGNMENTS

```

RESULT 1
ID MYHD_HUMAN STANDARD: PRT: 1938 AA.
AC 09UKX3: 095252:
DF 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN Myosin heavy chain, skeletal muscle, extraocular (MYHC-ec).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Extraocular muscle;
RX MEDLINE=99318869; PubMed=10388558;
RA Weiss A., Schiaffino S., Leinwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
RT heavy chain family: implications for functional diversity.";
RL J. Mol. Biol. 290:61-75(1999).
RN [2]
RP SEQUENCE OF 1917-1938 FROM N.A.
RC TISSUE=Extraocular muscle;
RX MEDLINE=99026150; PubMed=9806854;
RA Winters L.M., Briggs M.M., Schachar F.;
RT "The human extraocular muscle myosin heavy chain gene (MYH13) maps to
RT the cluster of fast and developmental myosin genes on chromosome 17.";
RL Genomics 54:188-189(1998).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MERKOTOSIN (LMH) AND 1 HEAVY MERKOTOSIN (HMH). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IO DOMAIN.
CC
CC *****
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC *****
CC EMBL: AF111782; AAC29948.1; -
CC DR EMBL: AF075248; AAC83241.1; -

```

DR HSSP: P08799; 1MND.
 DR MIM: 603487; -.
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 DR Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KM Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KM Multigene family.
 KM DOMAIN 1 784 MYOSIN HEAD-LIKE.
 FT DOMAIN 1 785 IQ.
 FT DOMAIN 843 1938 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP (POTENTIAL).
 FT DOMAIN 659 681 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 761 775 ACTIN-BINDING (TR1-) (POTENTIAL).
 FT MOD_RES 130 130 METHYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 699 699 ALKYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 709 709 ALKYLATION (SH-2) (POTENTIAL).
 SQ SEQUENCE 1938 AA; 223678 MW; 1F6D006416381CDS CRC64;

Query Match 8.1%; Score 192.5; DB 1; Length 1938;
 Best Local Similarity 20.9%; Pred. No. 0.018;
 Matches 94; Conservative 74; Mismatches 161; Indels 121; Gaps 16;

QY 3 IRAICTCSDFDH--SRDVAIHGHTPHLCILOSFEAPSTCCQCRIVQKRTIIN 60
 DB 1249 IERTCTVEQFSEIKAKDQQTQLIHLDNKO-----KARLQTONGELSH 1293
 QY 61 KLFDPLAEEENVL-----REFIKNELDNVRAQLSQDKK-----RDSQVITDILR 108
 DB 1294 RV-----EKRESLSQLTGSKKQALTOLEELKRMEEETAKNMAMALOSSRDCCLLR 1348
 QY 109 DYLEENATVVSLOALGKAE-----MICSTLKKOMKYLEQOODETRQAO 153
 DB 1349 EOYEEDDEAKAEIQRALSKANSEVAQWKTYETDAIQRTLELEBAKKKLAQRLQEAEREK 1408
 QY 154 EELGRLLSKKKTMEIOILLLOSLPEVEEMIRDMGVGQSAVEQLAVVCSLAKREYNLK 212
 DB 1409 ETA---NSKASLEKTKORLQ---EVEDLMRD-----ERSHTACATLDKKORNFDK 1455
 QY 213 -----BARKASGEVADK---LRKDLFSSRSKLTQTVYSILDQAKLELSAQKDLOS 259
 DB 1456 VLAEMWKQKLDSEQAELEBAQKRSKSLSTELFKMNAYEEVVDQLETLRREKKNQGEISD 1515
 QY 260 ADEEINSLKKKLMQET-----LNLPPASE 286
 DB 1516 LTRQIAETGNLQEAETKRLVPEQKSDLOVALEEVGSLSEHESKILRVQLSELQYKSE 1575
 QY 287 TVDLRLVESAPAEVNVNKLRRPSFRDDIDNATFDVD---TPRPASSSGHGYYETL 342
 DB 1576 -LDRKYLEKEDEIE---QLKRNQORAAEALQSVLDAEIRSNDAIRLKKMEGDLNMEI 1631
 QY 343 EKSHSPIQDVPK---KICKGPRKESQLSL 368
 DB 1632 QLGHSNRQMAETQRHLRTVQGLKDSQLHL 1661

RESULT 2
 CENE_HUMAN STANDARD; PRT; 2663 AA.
 AC 002234;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update).
 DE Centromeric protein E (CENP-E protein).
 GN CENPE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93024922; PubMed=1406971;
 RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.;
 RT "CENP-E is a putative kinetochore motor that accumulates just before
 RT mitosis.";
 RL Nature 359:536-539(1992).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=95196755; PubMed=7889940;
 RA Thirmer D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.;
 RT Mitotic HeLa cells contain a CENP-E-associated minus end-directed
 RT microtubule motor.";
 RL EMBO J. 14:918-926(1995).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=98437347; PubMed=9763420;
 RA Chan G.K.T., Schaar B.T., Yen T.J.;
 RT "Characterization of the kinetochore binding domain of CENP-E reveals
 RT interactions with the kinetochore proteins CENP-F and hNubrl.";
 RL J. Cell Biol. 143:49-63(1998).
 CC -I- FUNCTION: MINUS-END DIRECTED MICROTUBULE MOTOR. PROBABLE
 CC KINETOCHORE MOTOR. ACCUMULATES JUST BEFORE MITOSIS AT THE G2 PHASE
 CC OF THE CELL CYCLE. PROBABLY IMPORTANT FOR CHROMOSOME MOVEMENT
 CC AND/OR SPINDLE ELONGATION.
 CC -I- SUBUNIT: INTERACTS WITH CENP-F AND BUBR1 KINASE.
 CC -I- SUBCELLULAR LOCATION: ASSOCIATES WITH KINETOCHORES DURING
 CC CONGRESSION, RELOCATES TO THE SPINDLE MIDZONE AT ANAPHASE. AND IS
 CC QUANTITATIVELY DISCARDED AT THE END OF THE CELL DIVISION.
 CC -I- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z15005; CA78727.1; -.
 DR PIR: S28261; S28261.
 DR HSSP: P17119; 3KAR.
 DR MIM: 117143; -.
 DR InterPro: IPR001752; kinesin.
 DR Pfam: PF00225; kinesin; 1.
 DR PRINTS: PR00380; KINESINHEAVY.
 DR SMART: SM00129; KISC; 1.
 DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
 KM Motor protein; Cell division; ATP-binding; Coiled coil; Mitosis;
 KM Cell cycle; Centromere.
 FT DOMAIN 1 335 KINESIN-MOTOR.
 FT DOMAIN 336 2471 COILED COIL (POTENTIAL).
 FT NP_BIND 86 93 GLOBULAR (POTENTIAL).
 FT NP_BIND 86 93 ATP (BY SIMILARITY).
 SQ SEQUENCE 2663 AA; 312087 MW; CFC13880C8C8CB8 CRC64;

Query Match 7.9%; Score 187.5; DB 1; Length 2663;
 Best Local Similarity 25.1%; Pred. No. 0.045;
 Matches 93; Conservative 66; Mismatches 133; Indels 79; Gaps 18;

QY 55 KRTIINKLFDPLAEEENVLDRREFIKNELDNVRAQLSQDKK--RDSQVITDILRPTLEE 113
 DB 1632 KMTAVN-----ETQKCKEIEHLKEQFETQKLNLENITENIRLTOI---LHENLEP 1680


```

RL J. Cell Biol. 105:913-925(1987).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEAVYMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.
CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMAIC ACTIVITY.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 IO DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00624; CAA68663.1; -.
DR PIR: A27224; A27224.
DR HSSP: P08799; 1MND.
DR InterPro: IPR000048; IO.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IO; 2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IO; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IO; 1.
KW Myosin; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
KW Methylation; Alkylation; Phosphorylation; Multigene family.
FT DOMAIN 1 789 MYOSIN HEAD-LIKE.
FT 790 819 IO.
FT 848 1509 COILED COIL (POTENTIAL).
FT DOMAIN 848 1226 ALPHA-HELICAL TAILPIECE (S2).
FT DOMAIN 1227 1252 HINGE.
FT DOMAIN 1253 1509 LIGHT MEROMYOSIN (LMM).
FT DOMAIN 1253 1482 ALPHA-HELICAL TAILPIECE (LMM).
FT DOMAIN 1483 1509 NONHELICAL TAILPIECE.
FT NP_BIND 182 189 ATP.
FT 660 682 ACTIN-BINDING.
FT DOMAIN 766 780 ACTIN-BINDING.
FT MOD_RES 133 133 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 700 700 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 1489 1489 PHOSPHORYLATION.
FT MOD_RES 1494 1494 PHOSPHORYLATION.
FT MOD_RES 1499 1499 PHOSPHORYLATION.
SQ SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;

```

```

Query Match 7.4%; Score 176; DB 1; Length 1509;
Best Local Similarity 23.4%; Pred. No. 0.088;
Matches 97; Conservative 74; Mismatches 156; Indels 88; Gaps 13;

```

60 NKLFFDLAQQEENYLDRELFKNELDNVRAOJSOKD-----KEKRSD 101

DB 894 DKLEKDLALALKLIDEGEKADLEEDNALQKVAAGLEELQETSANDILQKRLE 953

QY 102 VIIDRLDLEERNATVVSLOALGKAE-----MLCSLTKOMKMYLEQOD 147

DB 954 AEKGLKASLDEEERNRAADQAKTVSESENNELQDKEDDAAHADSLKKKEEDSLRLR 1013

QY 148 ETKQAQEBAGR---LRSKMKTMQEIILLLOSLPEV-EEIMRDMGVGSAAVEOLAVYCV 202

DB 1014 ETKDALADANISETLKSLKLNTERGADVNRNEIDVTATKLOLEKTKKSLSEELQTRA 1073

QY 203 SLKREYNLEKARKASGEVADKLDKDFSSRSKQIQTYYSELDQAKLEKSAQ---KDLQ- 258

DB 1074 QLEBE-----KSGKFAASAKKQLOQOULEDARSEVDSLKSLSAEKSLEKTRAKDONRDLDE 1129

QY 259 -----SADKEINSLKRLKLTMLQETL-----NLPPVASETVDRVLSPAPE- 300

DB 1130 QLEDEFTVANVDKQKALEKLTLELDQYVALDGKNNAQAQKLTQVDEKRRLEE 1189

QY 301 -----VNLKLRBSEFRDDIDNATFDVDTTPARPSSQHGVEKLCLEKSHPLQDVPK 355

DB 1190 AEASARLEKERNKAIDEV---ADLTADLAERDSGAQORRKILNTRISELQSELENAVK- 1245

QY 356 ICKPFRRESQSLISGGSCAGPEDEIVGAPPIFRNAILQOKOPKRRPSSSCSK 410

DB 1246 -TGGASSEVYKRLGE--LERLBEELIYA-----QBARAAEKNIIDK 1284

RESULT 7

MYSS_CYPCA STANDARD; PRT; 1935 AA.

ID MYSS_CYPCA

AC Q90339;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin heavy chain, fast skeletal muscle.

OS Cyprinus carpio (Common carp).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;

OC Cypriniformes; Cyprinidae; Cyprinus.

OX NCBI_Taxid=7962;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Fast muscle;

RX MEDLINE=97352533; PubMed=9208928;

RA Hirayama Y., Watabe S.;

RT "Structural differences in the crossbridge head of temperature-

RT associated myosin subfragment-1 isoforms from carp fast skeletal

RT muscle.";

RL Eur. J. Biochem. 246:380-387(1997).

RN [2]

RP SEQUENCE OF 981-1935 FROM N.A.

RC TISSUE=Fast muscle;

RX MEDLINE=97176447; PubMed=9023993;

RA Imai J., Hirayama Y., Kikuchi K., Kakinuma M., Watabe S.;

RT "cDNA cloning of myosin heavy chain isoforms from carp fast skeletal

RT muscle and their gene expression associated with temperature

RT acclimation.";

RL J. Exp. Biol. 200:27-34(1997).

RN [3]

RP SEQUENCE OF 1387-1528 FROM N.A.

RX MEDLINE=95194396; PubMed=787920;

RA Watabe S., Imai J., Nakaya M., Hirayama Y., Okamoto Y., Masaki H.,

RA Uozumi T., Hirose I., Aoki T.;

RT "Temperature acclimation induces light meromyosin isoforms with

RT different primary structures in carp fast skeletal muscle.";

RL Biochem. Biophys. Res. Commun. 208:118-125(1995).

CC -1- FUNCTION: MUSCLE CONTRACTION.

CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2

CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)

CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).

CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.

CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING

CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,

CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY

CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.

CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT

CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE

CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED

CC SUBFRAGMENT (S2).

CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D89992; BAA22069.1; -;
 DR EMBL: D50476; BAA09069.1; -;
 DR EMBL: D43700; BAA07802.1; -;
 DR HSSP: P08799; 1MMD.
 DR InterPro: IPR000048; IO.
 DR InterPro: IPR000009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IO; 1.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IO; 1.
 KW Myosin; muscle protein; coiled coil; thick filament; actin-binding;
 KW ATP-binding; Methylation; Alkylation; Calmodulin-binding;
 KW Multigene family.
 FT DOMAIN 1 781 MYOSIN HEAD-LIKE.
 FT DOMAIN 782 811 HINGE.
 FT DOMAIN 812 839 COILED COIL (POTENTIAL).
 FT DOMAIN 840 1935 ATP (POTENTIAL).
 FT NP_BIND 178 185 ACTIN-BINDING.
 FT DOMAIN 659 681 ACTIN-BINDING.
 FT DOMAIN 761 775 METHYLATION (TRI-) (POTENTIAL).
 FT MOD_RES 129 129 METHYLATION (SH-1).
 FT MOD_RES 699 699 ALKYLATION (SH-1).
 FT MOD_RES 709 709 ALKYLATION (SH-2).
 SO SEQUENCE 1935 AA; 221599 MW; 9A1244B67D63C83B CRC64;

Query Match 7.3%; Score 173.5; DB 1; Length 1935;
 Best Local Similarity 22.3%; Pred. No. 0.16;
 Matches 72; Conservative 66; Mismatches 126; Indels 59; Gaps 11;

QY 31 LQCLIQSETPAPSRCTPCQRCIQVGRKRTITINKLFFDLAODEENVLDRF----- 78
 DB 1144 LEETISERLEEGAGTAAQIEENKREAEFOKMRDL--EESTLQHEXTAALRKEQADS 1200
 QY 79 ---LKNELDNVRAQLSQDKDEKRDQVYIDTLRPTLEERNNTVVSIDQALQAKMLCSTL 135
 DB 1201 VAEIGEDIDNQVRKQKLEKESYKMEIDLTJNME-----AVAKAKNLEKCKRTL 1253
 QY 136 KKKQKYLEQOQDETKQAOEAGRLSRKMT-----MEQIELLLQSOLP----- 178
 DB 1254 EDQISEIKTKSDENVRLQNDNNARARLQTFENGFEFSRLQLEKALV--SQLTRGQATYQQ 1312
 QY 179 -----EVEEMIRDMGVGQSAVEQDLAVYCVSLKREYENLKKARRASGEVADKLKDLFSS 232
 DB 1313 IEELKRRIIEEVRKAKNALAHVQSAHRDCLDLRQYEEEOGRKAKAELRGMSKANSEVAQW 1372
 QY 233 RSKLQI-----VYSLDQAKLELKSQKDLQADKSEIMSLKKLMLQETLNLPPVASETV 288
 DB 1373 RTKEETAIQRTLEELKPKKL--AQR-LDAEESIEAVNSKASLEKTYQ--RLQGEVE 1427
 QY 289 DRV-LLESPAPVEYNLKRPSF 310
 DB 1428 DLMIDVERANSIANTLKKQKNF 1450

RESULT 8
 MYH4_RABBIT STANDARD; PRT; 1938 AA.
 AC Q2864;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, skeletal muscle, juvenile.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP STRAIN=NEW ZEALAND WHITE; TISSUE=Skeletal muscle;
 RC Mada K., Hostilnova E., Roesch-Kleinhauf A., Schuster H., Gasperik J.,
 RA Wittlinghofer A.;
 RT "Isolation, sequencing of myosin heavy chain cDNA from rabbit
 RT skeletal muscle and a novel cosynthesis of S-1 fragment with the
 RT essential and regulatory light chains";
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBUNIT: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U32574; AAA74199.1; -;
 DR HSSP: P08799; 1MMD.
 DR InterPro: IPR000048; IO.
 DR InterPro: IPR004008; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IO; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IO; 1.
 KW Myosin; muscle protein; coiled coil; thick filament; actin-binding;
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KW Multigene family.
 FT DOMAIN 1 783 MYOSIN HEAD-LIKE.
 FT DOMAIN 784 813 IO.
 FT DOMAIN 812 1938 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP (POTENTIAL).
 FT DOMAIN 658 680 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 774 774 ACTIN-BINDING (BY SIMILARITY).
 FT MOD_RES 35 35 METHYLATION (MONO-) (BY SIMILARITY).
 FT MOD_RES 130 130 METHYLATION (TRI-) (BY SIMILARITY).
 FT MOD_RES 552 552 METHYLATION (TRI-) (BY SIMILARITY).
 FT MOD_RES 756 756 METHYLATION (SH-1) (BY SIMILARITY).
 FT MOD_RES 698 698 ALKYLATION (SH-1) (BY SIMILARITY).

FT MOD.RES 708 708 ALKYLATION (SH-2) (BY SIMILARITY).
SQ SEQUENCE 1938 AA: 223064 MW: D8A8A2EC5B182626 CRC64;
Query Match 7.2%; Score 172; DB 1; Length 1938;
Best Local Similarity 21.3%; Pred. No. 0.18;
Matches 86; Conservative 82; Mismatches 148; Indels 88; Gaps 16;
OY 31 LOCILISFETAPSPRCRIGVGRRTITINKLFEDLAQEEENVALREF----- 78
DB 1146 LEISISRLLEAGATSAQILEMKNKKRAEFCKMRDL---EATLQHEATATATIRKKHADS 1202
OY 79 ---LKNELDNVRAQLSQDKREKRDQVIITDLRDLTEERNATVVSLOALGKAEMLCSTL 135
DB 1203 VAEIGBOIDNLQVOKOLEEKSELMKEIDDLASNME-----TVSKAKNGLEKKMCRIL 1255
OY 136 KQOMKTLBOQOQDETK-----QAQEEAGRL-----RSKMKTMEOI 169
DB 1256 EDQVSELKTRKEEHQRLINDLSAQARARLQTESEGFSSROLDKNDLSVQLSRGQAFQOI 1315
OY 170 ELLLOSQLPEVEEMIRDMGVQSAVEQLAVYCVSLKKEVENLKERRASGEVADKLRLKD 229
DB 1316 EEELK-----QLEEFIRAKSLAHALOSARHDCDLLREQTEEOEAKAELORAMSKANSEV 1371
OY 230 FSSRSKLQT---VYSELQDAKLELSAOKDQSADEKELMSLKKULMLQETLNPVVAS 285
DB 1372 AQGRTEYEDAIQTELELEBAKKL--AQR-LQDAEHEVAVNAKASLEKTKQ--RLQN 1426
OY 286 EYVDRLV-LESAPAVEVNLKLRPSFRDI-----DLNATFVDVTPPARPSSQ-- 333
DB 1427 EYEDLMDIADERTINACALADKQRNF-DKLIAMKHYETHAELFASOKESRSLSLEVP 1485
OY 334 ---HGVEKL---CLEKSHSPIQ---DVKKKICKGRKESOL 366
DB 1486 KVNNAIEESLDQLETLKREKNKLQOEISDTEQIABGKRIHEL 1529
RESULT 9
MYS2.DICDI
ID MYS2.DICDI STANDARD; PRT; 2116 AA.
AC P08799;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MYOSIN II heavy chain, non muscle.
GN MHCA.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=87092266; PubMed=3540939;
RA Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;
RT "Conserved protein domains in a myosin heavy chain gene from
RT Dictyostellium discoideum.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN (2)
RP PHOSPHORYLATION SITES, AND MUTAGENESIS.
RC STRAIN=AX2;
RX MEDLINE=90353583; PubMed=2387408;
RA Lueck-Vielmeier D., Schleicher M., Grabatin B., Wippler J.,
RA Gerisch G.;
RT "Replacement of threonine residues by serine and alanine in a
RT phosphorlatable heavy chain fragment of Dictyostellium myosin II.";
RL FEBS Lett. 269:239-243(1990).
RN (3)
RP PHOSPHORYLATION SITES.
RX MEDLINE=88112226; PubMed=2828113;
RA Wagle G., Noegel A., Scheel J., Gerisch G.;
RT "Phosphorylation of threonine residues on cloned fragments of the
RT Dictyostellium myosin heavy chain.";
RL FEBS Lett. 227:71-75(1986).
RN

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RX MEDLINE=95345066; PubMed=7619795;
RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
RA Rayment I.;
RT "X-ray structures of the myosin motor domain of Dictyostellium
RT discoideum complexed with MgADP.Befx and MgADP.Alfa-";
RL Biochemistry 34:8960-8972(1995).
RN (5)
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RX MEDLINE=95345067; PubMed=7619796;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the
RT truncated head of Dictyostellium discoideum myosin to 2.7-A
RT resolution.";
RL Biochemistry 34:8973-8981(1995).
RN (6)
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX MEDLINE=96206189; PubMed=8611530;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the
RT Dictyostellium discoideum myosin motor domain to 1.9-A resolution.";
RL Biochemistry 35:5404-5417(1996).
RN (7)
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
RX MEDLINE=97452580; PubMed=9305951;
RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RT "X-ray structures of the MgADP, MgATPamas, and MgADPPNP complexes
RT of the Dictyostellium discoideum myosin motor domain.";
RL Biochemistry 36:11619-11628(1997).
RN (8)
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX MEDLINE=98070605; PubMed=9405148;
RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
RT "X-ray crystal structure and solution fluorescence characterization
RT of Mg.2(3')-O-(N-methylanthraniloyl) nucleotides bound to the
RT Dictyostellium discoideum myosin motor domain.";
RL J. Mol. Biol. 274:394-407(1997).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY ACTIN.
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC (MLC-2).
CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
CC COREX.
CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN BE FURTHER
CC SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
CC THE ACTIN-ACTIVATED ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: DICTYOSTELLUM MYOSIN II HAS NO K(2)EDTA ATPASE
CC ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
CC POSITION (688).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 TO DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M14628; AAA33227.1; -
DR PIR: A26655; A26655.
DR PIR: S00250; S00250.
DR PDB: 1MMA; 03-DEC-97.
DR PDB: 1MMD; 17-AUG-96.
DR

DR PDB: 1MKG: 03-DEC-97.
 DR PDB: 1MNN: 03-DEC-97.
 DR PDB: 1MND: 17-AUG-96.
 DR PDB: 1MNE: 17-AUG-96.
 DR PDB: 1VOM: 23-DEC-96.
 DR PDB: 1LVK: 28-JAN-98.
 DR DictyDb: DD01008: mhca.
 DR InterPro: IPR000048: IQ.
 DR InterPro: IPR004009: Myosin_N.
 DR InterPro: IPR001609: myosin_head.
 DR Pfam: PF00612: IQ: 2.
 DR Pfam: PF00063: myosin_head: 1.
 DR Pfam: PF02736: Myosin_N: 1.
 DR PRINTS: PR00193: MYOSINHEAVY.
 DR ProDom: PD000355: myosin_head: 1.
 DR SMART: SM00015: IQ: 1.
 DR SMART: SM00242: MYSC: 1.
 DR PROSITE: PS50096: IQ: 1.
 KM Myosin: Coiled coil: Actin-binding; ATP-binding; 3D-structure;
 KM Calmodulin-binding; Methylation; Alkylation; Phosphorylation.
 FT DOMAIN 1 761 MYOSIN HEAD-LIKE.
 FT DOMAIN 762 791 IQ.
 FT NP_BIND 817 2116 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP.
 FT DOMAIN 638 660 ACTIN-BINDING.
 FT DOMAIN 738 752 ACTIN-BINDING.
 FT MOD_RES 130 130 METHYLATION (DI-) (POTENTIAL).
 FT MOD_RES 678 678 ALKYLATION (SH-1).
 FT MOD_RES 1823 1823 PHOSPHORYLATION (BY MHCK).
 FT MOD_RES 1833 1833 PHOSPHORYLATION (BY MHCK).
 FT MOD_RES 2029 2029 PHOSPHORYLATION (BY MHCK).
 SQ SEQUENCE 2116 AA: 243871 MW: 2FC3770BB1E56A1 CRC64;

Query Match 7.2%; Score 172; DB 1; Length 2116;
 Best Local Similarity 27.4%; Pred. No. 0.2;
 Matches 90; Conservative 49; Mismatches 109; Indels 80; Gaps 17;

QY 60 NKLFPDLAQQEENVLD-REFLNKELNDVNRALQSKD---KEKRDQVITITLDTLEERN 115
 DB 852 DKLEKSLKQTESNVLDRQLKAEKELTKAMYSKDALEKQRELETRVEDMESELEDEKK 911
 QY 116 ATTVSLQALGKAEMLCTLKQMKYLEQOODETKQAQAEGRLSKMKTEOIELLOS 175
 DB 912 LALENLQNKRSVE-----EKVRDELEELQEE-----QKLRNLEKIKKK 951
 QY 176 QLPEVEEMIRDMGVGS-AVEQLAVYCVSLKKEYENLEKARKASGEVADK--LRKDLFSS 232
 DB 952 YEELLEEMKR-VNDGSDTISRLEKIKDELQKEVEELTES--FSEESKDKGVLEK---T 1004
 QY 233 RSKLQTVYSLDQAKLELTKSAOKDLSADKEIMSLKKKKLTMLQETLNLPPVASFTVDRLV 292
 DB 1005 RVRLO--SELDLTVNLDETQKSELRLQKKKLEELQVOQAL-----AAETAACKLA 1056
 QY 293 LESPAPEVNLKLRPSFRDIDLNTATFDVTPPARSSQHGYYEKLCEKSHSPIDV 352
 DB 1057 QEA-----ANKKQ-----GGYTELENEKFNSEV--TARSN-----VEKSKTLL--- 1092
 QY 353 PKKICKPKRESQSLSGQSCAGEPDEE 380
 DB 1093 -----ESQL-----VAVNNELDEE 1106

RESULT 10
 ID RA50_AQUAE STANDARD; PRT; 978 AA.
 AC 067124;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Probable DNA double-strand break repair rad50 ATPase.
 GN RAD50 OR AQ_1006.
 OS Aquifex aeolicus.

OC Bacteria: Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjaj M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 392:353-358(1998).
 CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
 CC rad50/mre11 complex possesses single-strand endonuclease activity
 CC and ATP-dependent double-strand-specific exonuclease activity.
 CC Rad50 provides an ATP-dependent control of mre11 by unwinding
 CC and/or repositioning DNA ends into the mre11 active site (By
 CC similarity).
 CC -1- SUBUNIT: Forms a complex with mre11 (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).
 CC -----
 DR EMBL: AE000718; AAC07092.1; -
 DR InterPro: IPR003439: ABC_transport.
 KM DNA repair: Hydrolyase; ATP-binding; Coiled coil; Complete proteome.
 FT NP_BIND 32 39 ATP (BY SIMILARITY).
 FT DOMAIN 160 826 COILED COIL (POTENTIAL).
 SQ SEQUENCE 978 AA: 115897 MW: 9B0F2BF51ADD1151 CRC64;

Query Match 7.1%; Score 170; DB 1; Length 978;
 Best Local Similarity 22.6%; Pred. No. 0.11;
 Matches 79; Conservative 70; Mismatches 125; Indels 76; Gaps 13;

QY 41 APSKTCPOCRITQVGRITIN-----KLFPDLAQQEENVLDREF-----LKNE 82
 DB 481 SPGDYCPGCGIRGKALEVNDAGEISELKHAKELKEKEIREIDTTLKVAQKINSKEE 540
 QY 83 LDVNRALQSKQKREKRDQVITITLDRDTEERNATVYSLQALGKAEMLCTLKQMKYL 142
 DB 541 MEKLRNEVEELRKE-----IPENLKER---IKLELEIRIEKEKLEHKLNKRYKKAL 587
 QY 143 EQOODETKQAQAEGRLSKMKTEOIELLOSQ--PEVEEMIRDMGVGSAYEQLAVYC 201
 DB 588 EDKQKQKEEQAOKLHKAQOTLELELKE-KIREKSLVKEFEELR-----VERLDYE 638
 QY 202 VSLKKEYENLEKARKASGEVADLRK--DLFSSRSKLTQVYSELDOAKLELTKSAOKDLO 258
 DB 639 ESLKEELINYSKLTQLELEKEKRLKHFELSSRKSLGEE-----ELALNESIN 688
 QY 259 SADKEIMSLKKTLMQETLNLPPVASFTVDRLVLESAPVE-----VNLKLRP 308
 DB 689 SLEEKREKLELANIYEVAKSP---REVVELYLGKQEALEKRIKFEESFOSLKLKS 745
 QY 309 SFRDDIDLNTATFDVTPPARSSQHGYYE--KLCEKSHSPIDVPKTI 356
 DB 746 EIEEKLK-----EYEGIRELSDIRGEYSVKTQLEKHKLGVEYKRL 788

RESULT 11
 ID TANA_XENLA STANDARD; PRT; 1744 AA.
 AC 001550;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)


```

Query Match          7.1%; Score 169.5; DB 1; Length 1940;
Best Local Similarity 23.0%; Pred. No. 0.25; Mismatches 110; Indels 65; Gaps 10;
Matches 70; Conservative 59;

OY 15 DHSRDVAIHCGHGFHLOCLIOSEFTAPSRTPCPCRIQVGRRTIINKLFEDLAQEEENVL 74
D 1139 DYARE-----LEELSEERLEAGGVSTQIELNKKREAEFLKRLDRL---EEATL 1184
OY 75 DREF-----LKNELDNVAQLSQDKENRDSQVIIDPLRLLEERNATV 119
D 1185 QHEATVATLRKKHADSAAELAQIDNLRQVKQLEKESEKLELDLSSSVE----- 1237
OY 120 SLOQALGKAEMICSTLKQKMYLYEODDETRKA-----QEEAGRISRKMTKEQ 168
D 1238 SVSKSKANLEKICRTLEQSLSEARGKNNEIORSISELTQKSRQTEGELSRQLEEKES 1297
OY 169 IELL---OSQDPEVEEMTRDMGVGOSAVBOLAVY-----CVSLKKEYENKEARKA 217
D 1298 IVSOLSSSKQAFQTOQIEELKQLEENKAKNALAHALQSRHDCDLLREQYEEDEQKAE 1357
OY 218 SGEVADKRLKDLFSSRSKLOF---VSELDQAKLELKSADKQDSADKELMSLKKILTM 273
D 1358 LQRLALSKANSEVQAWRKRYETDAIQRTLEELKAKKL--AQR-LQDSEQVEAVNAKCAAS 1414
OY 274 LQET 277
D 1415 LEKT 1418

RESULT 13
MYH7_RAT
ID MYH7_RAT STANDARD; PRT; 1935 AA.
AC P02564;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, cardiac muscle beta isoform (Myhc-beta).
GN MYH7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=80016823; PubMed=2798112;
RA Kraft R., Bravo-Zehnder M., Taylor D., Leinwand L.A.;
RT "Complete nucleotide sequence of full length cDNA for rat beta
RT cardiac myosin heavy chain."
RL Nucleic Acids Res. 17:7529-7530(1989).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=90133919; PubMed=2614840;
RA McNally E.M., Kraft R., Bravo-Zehnder M., Taylor D., Leinwand L.A.;
RT "Full-length rat alpha and beta cardiac myosin heavy chain sequences.
RT Comparisons suggest a molecular basis for functional differences.";
RL J. Mol. Biol. 210:665-671(1989).
RN [3]
RP SEQUENCE OF 1524-1935 FROM N.A.
RX MEDLINE=82220036; PubMed=7045682;
RA Mahdavi V., Perlasamy M., Nadal-Ginard B.;
RT "Molecular characterization of two myosin heavy chain genes expressed
RT in the adult heart."
RL Nature 297:659-664(1982).
RN [4]
RP SEQUENCE OF 1871-1935 FROM N.A.
RC STRAIN=MISTAR; TISSUE=Heart;
RX MEDLINE=85179510; PubMed=6241892;
RA Mahdavi V., Lompre A.M., Chambers A.P., Nadal-Ginard B.;
RT "Cardiac myosin heavy chain isozyemic transitions during development
RT and under pathological conditions are regulated at the level of mRNA
RT availability."

```

```

RL Eur. Heart J. 5:181-191(1984).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- MISCELLANEOUS: THE CARDIAC ALPHA ISOFORM IS A 'FAST' ATPASE
CC MYOSIN, WHILE THE BETA ISOFORM IS A 'SLOW' ATPASE.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X15939; CA34065.1; -.
CC EMBL: J00752; AAA41654.1; -.
CC EMBL: M32698; AAA41659.1; -.
CC PIR: S06006; S06006.
CC PIR: A02989; A02989.
CC HSP: P08799; 1MMD.
CC InterPro: IPR000048; IQ.
CC InterPro: IPR004009; Myosin_N.
CC InterPro: IPR002928; Myosin_tail.
CC InterPro: IPR001609; myosin_head.
CC Pfam: PF00612; IQ_2.
CC Pfam: PR00063; myosin_head; 1.
CC Pfam: PF02736; Myosin_N; 1.
CC Pfam: PF01576; Myosin_tail; 1.
CC PRINTS: PR00193; MYOSINHEAVY.
CC ProDom: PD000355; MYOSINHEAD; 1.
CC SMART: SM00015; IQ; 1.
CC SMART: SM00242; MYSC; 1.
CC PROSITE: PS50096; IQ; 1.
CC MYOSIN: Muscle protein; Coiled coil; Thick filament; Actin-binding;
CC ATP-binding; Methylation; Alkylation; Multigene family;
CC Calmodulin-binding.
CC DOMAIN 1 780
CC FT 781 810 MYOSIN HEAD-LIKE.
CC FT 810 810 IQ.
CC FT 810 810 COILED COIL (POTENTIAL).
CC FT NP_BIND 178 185 ATP.
CC FT DOMAIN 655 677 ACTIN-BINDING.
CC FT DOMAIN 757 771 ACTIN-BINDING.
CC FT MOD_RES 129 129 METHYLATION (TRI-1) (POTENTIAL).
CC FT MOD_RES 695 695 METHYLATION (SH-1) (POTENTIAL).
CC FT MOD_RES 705 705 ALKYLATION (SH-2) (POTENTIAL).
CC FT CONFLICT 1529 1531 ALKYLATION (VRR (IN REF. 3)).
CC FT CONFLICT 1731 1731 D -> H (IN REF. 3).
CC FT CONFLICT 1784 1784 N -> K (IN REF. 3).
CC FT CONFLICT 1851 1851 T -> N (IN REF. 3).
CC FT CONFLICT 1858 1858 R -> K (IN REF. 3).
CC SEQUENCE. 1935 AA; 223082 MW; C8376C32A47BD82B CRC64;

```

```

Query Match          7.1%; Score 169; DB 1; Length 1935;
Best Local Similarity 22.5%; Pred. No. 0.26; Mismatches 106; Indels 56; Gaps 9;
Matches 65; Conservative 62;

OY 31 LOCLIOSEFTAPSRTPCPCRIQVGRRTIINKLFEDLAQEEENVLDREF----- 78
I : : : | | | : : : | : | | | |

```

[illegible]

```

FT DOMAIN <1 619 GLOBULAR HEAD
FT DOMAIN 620 912 COILED COIL (POTENTIAL).
FT DOMAIN 917 941 ASP/GLU-RICH (ACIDIC).
FT MOD_RES 922 922 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 941 AA; 105152 MW; AB5C10895CD7E508 CRC64;

Query Match 7.0%; Score 168; DB 1; Length 941;
Best Local Similarity 19.6%; Pred. No. 0.13;
Matches 74; Conservative 55; Mismatches 96; Indels 152; Gaps 10;

QY 52 QVGRKTIINKLFF-----DLAQEEENVLDREFLK--NELDNV--RAQDS 91
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 553 RIGKENFLEKLGFIKSNHELYSRASQKPPNPFPSPYMI FDEHFTLVKLEGVITKAIYK 612
QY 92 QDKDRDSQVYIDTLRPTLEERNATVYSLOALKAEMICSTLKOKMKYLEEQODE--- 148
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 613 SSEEKKEEV-----KTLIEDHNVITHYKMMITEQDLQLEELKQOYSTLKQNEQLQ 667
QY 149 -TKQAQ-----EEAGRLSRKMTMQI 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 668 AVTQQAQSIQHQKDDYNLLKVLQDGKNNHQSGBGAQVNGIOPREISRLREIQLSKQ 727
QY 170 ELLDSQLEPEVEMTRDKGVGQSA--VEQLAVYC----- 201
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 728 QALLOGOLAEXDLSLEIENLKSSQASQMSQASATCPRPPEQVAVELKQELTALKSOLCSQS 787
QY 202 -----VSLKKEYENLEAKR 216
DB 788 LEITFLQENCCELLQRAETLAKSVPEGESEHVSAAKTTDVEGRSLALQETKELKNEIK 847
QY 217 ASGEVADLRKLDFFSRSKLTQTVSELDQAKLELKSQKD-----LOSADKEIIMSLRK 269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 848 ALSEERTAIQKQDSSNSTIATLQTEKPKLDLEVYDSKKEQDDLVLLADQOKITLSLKS 907
QY 270 KITMLQETLNLPPVASE 286
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 908 KKLKDLGH-----PVEEE 919

RESULT 15
G160_MOUSE STANDARD: PRT: 1325 AA.
AC P55937;
DT 01-NOV-1997 (Rel. 35, created)
DT 01-NOV-1997 (Rel. 35, last sequence update)
DT 15-DEC-1998 (Rel. 37, last annotation update)
DE Golgin-160 (Male-enhanced antigen-2) (Mea-2).
GN GOLGA3 OR MEA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Testis;
RX MEDLINE=97217683; PubMed=9063644;
RA Kondo M., Sutoh S.;
RT "Cloning and molecular characterization of cDNA encoding a mouse
RT male-enhanced antigen-2 (Mea-2): a putative family of the Golgi
RT autantigen."
RL DNA Seq. 7:71-82(1997).
CC - FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR
CC TESTIS DEVELOPMENT. PROBABLY IDENTICAL WITH THE SEROLOGICALLY
CC DETECTABLE MALE ANTIGEN (SDM).
CC - TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. TRANSCRIPTS CAN BE
CC FOUND IN SPERMATIDS DURING SPERMATOGENESIS. NO EXPRESSION IN
CC LEYDIG CELLS, SPERMATOZOONIA, OR SPERMATOCYTES.
CC - SIMILARITY: HIGH. NO HUMAN GOLGIN-160.
CC - CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-19 OR MET-30 IS THE
CC INITIATOR.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D78270; BAA19612.1; -
 DR HSSP: P18852; 1SCG.
 DR MGD: MGI:96958; Gciga3.
 KM Spermatogenesis; Developmental protein.
 FT DOMAIN 201 204 POLY-ALA.
 SQ SEQUENCE 1325 AA; 149880 MW; 3230636962C687B0 CRC64;

Query Match 7.0%; Score 168; DB 1; Length 1325;
 Best Local Similarity 20.6%; Pred. No. 0.19;
 Matches 77; Conservative 65; Mismatches 130; Indels 102; Gaps 13;

QY 68 QEEENVLDREFLNKELD-----NVRQOLSQDKERKDSQYIIDLRLDTEERNATVYSL 121
 DB 966 REHNSILETALAKREADVQNLQVAVLQKREEDRQMKQLVQALQVSLKEMEVNLSL 1025
 QY 122 QQALGKA-----EMLCSTLKRQMKYLEQO---ODET 149
 DB 1026 KEQMAAARIEAGNRRHFKATLELSEVKKELQAKELHLYOTLQAEVDELQIDQSKHSEI 1085
 QY 150 KQAEFAGRLSKKMTMEQIELLSQLEPEVEEMIRMGVQSAVEQALAVCVSLKKEYE 209
 DB 1086 AQOFTELAEATQQLQ-----LQKRL--DEQMSQOPTSQSMEDEKMKELDKERETIQ 1135
 QY 210 NLKEARRASGEVADKLRLKDLFFSSRSKQTVYSELQAKLELKSQKDLQSDAKREIMSLK 269
 DB 1136 SLKQQLDLTEQOG---KKELEGTOQTLOTIKSELEWQEDLSEFQKQKFMQLQAVSELKN 1192
 QY 270 KL-TMLOET-----LNLP--PVASETVDRVLVLESSPA 297
 DB 1193 NMKTLQDQNOQLKDLBRGAARKKEPKGESNSSSPATPIKIPDCPVASLLEELLRPPPA 1252
 QY 298 ----PVENVNLKLRPRFRDID-LNATEVDVTPPARPSSOHGYEKLCEKSSPIQDV 352
 DB 1253 VSKPEPLK-NLNNCLQQLQKQEWDSLQROMEEHTTVHESLSWAQVEAPAEHNAH----- 1305
 QY 353 PKKICKGPKRESQSL 366
 DB 1306 -----PRGDTKL 1312

RESULT 16
 MYH2_HUMAN STANDARD; PRT; 1941 AA.
 ID MYH2_HUMAN 016229; 014322;
 AC G9UKX2; 016229; 014322;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, skeletal muscle, adult 2 (Myosin heavy chain IIA)
 DE (MYHC-IIa).
 GN MYH2 OR MYHSA2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=99318869; Pubmed=10388558;
 RA Weiss A., Schiaffino S., Leinwand L.A.;
 RT "Comparative sequence analysis of the complete human sarcomeric myosin
 RT heavy chain family: implications for functional diversity.";
 RL J. Mol. Biol. 290:61-75(1999).
 RN [2]
 RP SEQUENCE OF 1711-1941 FROM N.A.
 RC TISSUE=Skeletal muscle;

RX MEDLINE=95109625; Pubmed=7545970;
 RA Shierdu V., Karsch-Mizrachi I., Campione M., Leinwand L.,
 RA Schiaffino S.;
 RT "type IIX myosin heavy chain transcripts are expressed in type IIB
 RT fibers of human skeletal muscle.";
 RL Am. J. Physiol. 267:C1723-C1728(1994).
 RN [3]

RP SEQUENCE OF 1823-1941 FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=95270723; Pubmed=7751403;
 RA Emilion S., Sant'ana Pereira J., Sargeant T., Young A., Goldspink G.;
 RT "Characterization of human skeletal muscle fibres according to the
 RT myosin heavy chains they express.";
 RL J. Muscle Res. Cell Motil. 16:35-43(1995).

CC -1- FUNCTION: MUSCLE CONTRACTION
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).

CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.

CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).

CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF11784; AAD29950.1; -
 DR EMBL: S73840; AAC13916.1; -
 DR EMBL: Z32856; CAAB3687.1; -
 DR HSSP: P08799; 1MND.
 DR MIM: 160740; -
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF02736; myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 KW Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KW Multigene family.
 FT DOMAIN 1 786 MYOSIN HEAD-LIKE.
 FT DOMAIN 787 816 IQ.
 FT DOMAIN 845 1941 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP (POTENTIAL).
 FT CONFLICT 1844 1844 K -> R (IN REF. 3)
 SQ SEQUENCE 1941 AA; 223043 MW; 681E866F83AEAB3F CRC64;

Query Match 7.0%; Score 168; DB 1; Length 1941;
 Best Local Similarity 22.6%; Pred. No. 0.29;
 Matches 66; Conservative 58; Mismatches 106; Indels 62; Gaps 9;

QY 31 LQCLIOSFETAPSRTPCQCRIVQGKRTIINKLFFDLAQEENVLDREF----- 78

```

Db 1149 LEEISERLIEBAGGATSAQIEMNKKREAEFOKMRDL---EEATLQHEATATLTKKKHADS 1205
QY 79 ---LKNELDNVRAQLOSKOKKEKPDISOVIITDLDTLEERNATVVSLOALGKEMLCSTL 135
Db 1206 VAEIGQIDMLQKQKLEKSEKSEKEMKIDDLASNVE-----TVSKAGNLEKMCRTL 1258
QY 136 KKGKMYLEQOQODETK-----QAQEEAGRL-----RSKMKTMEOI 169
Db 1259 EDLSELSKSEEBEQOQLINDLTQGRGLQTESESEFSDQDEKFAVLSQLSRGQAFQOI 1318
QY 170 ELLLOQLPVEVEEMIDMGVGSQVAVQLAVYCVSLKEKENLKEARKASGEVADKLKDL 229
Db 1319 EEIKR-----QLEEEIKAKNALAHALQSSRHDCDLREQYEEQESKAEIQLRALSKANTEV 1374
QY 230 FSSRSKQOT-----VYSLEQAKLELSAQKDLOSADKEINSLKKKLTMLDET 277
Db 1375 AQWRTYETDAIQRTLELEAKKRL--AQR-LQAEEHVAVNAKCASLKKT 1423

RESULT 17
MISP_ONCVO STANDARD; PRT; 879 AA.
ID MISP_ONCVO
AC 002171;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Paramyosin.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93165084; PubMed=7679477;
RA Dahmen A., Gallin M., Schumacher M., Ertmann K.D.;
RT "Molecular cloning and pre-mRNA maturation of Onchocerca volvulus
  paramyosin."
RL Mol. Biochem. Parasitol. 57:335-338(1993).
CC -1- FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF
  MANY THICK FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- SIMILARITY: HIGH, TO MYOSIN HEAVY CHAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL: M95813; AAA29431.1; -.
DR HSSP: P80220; IDIP.
DR InterPro: IPR002928; Myosin_tail.
DR Pfam: PF01576; Myosin_tail; 1.
KM Coiled coil; Muscle protein; Thick filament; Myosin.
FT DOMAIN 1 36 NONHELICAL REGION (POTENTIAL).
FT DOMAIN 37 860 COILED COIL (POTENTIAL).
FT DOMAIN 861 879 NONHELICAL REGION (POTENTIAL).
SQ SEQUENCE 879 AA; 101125 MW; DCA24AC01A9F02B1 CRC64;

Query Match 7.0%; Score 167; DB 1; Length 879;
Best Local Similarity 19.8%; Pred. No. 0.13;
Matches 105; Conservative 88; Mismatches 160; Indels 178; Gaps 20;

QY 59 INKLFEDLAE-----EEVLDREF--LKNELDN--VRAQLOSKOKKEK----- 98
Db 198 LNRHVNDLAAQQRQRLAENNDLKEITHQVQDDNLOHVAYQOLAQOOLEEARRLLEDAERE 257
QY 99 -----DSQVITIDTLRLDLEERNATVVSLOALGKA----- 128

```

```

Db 258 RSQOLAQLHOVQLELDVSFTALDEESARAARAEHKLALANTETITQKSKDAEVALHDEE 317
QY 129 --EMLCSTLKKQKKYLEQOO---DETKQAQDEEGRLSKMKRTY----- 166
Db 318 VEDLRKMKLOKAQAEIEMLOKISOLEKAKSRLOQSEVEVLIDLEKRAONTAILBERA 377
QY 167 -EOIEELLQSLPVEVEEMIDMGVGSQVAVQLAVYCVSLKEKENLKEARKASGEVADKLKDL 207
Db 378 KEQLEKTVNELKRIDELIVELVPAQREARALAELOKKKNLYEKAVEQKEALARENKKL 437
QY 208 YENLKEARKASGEVADKLK-DLFSSR-----SKLOTYVSEI 243
Db 438 QDQLHBAKEALDANRKLHELDENLRAGEIRELOQTALKESEPARDAENRQALAEI 497
QY 244 DQAKLELSAQKDLOSADKEINSLKKKLTMLQOTLMLPPASTVRL--VLESAPVE 300
Db 498 QQLRIEM--ERRLOKEEEMELRKNMF-----EIDRLAALADAEARKK 541
QY 301 VNLKRRPSFRDQI-DLNATFD-----VDTPPARPSSOHGYEKKLEKSHSPIQDVP 353
Db 542 AEIARLKRTYQAEIAELEMTVDNLNANIFAOKTIRKQSEQLKVLQASLEDTORLOQTL 601
QY 354 KIKCGPRKESQSLGQSCAGEPDELVGAFPIFYRNAILGQKQPKRPRESSCSKDVV 413
Db 602 DQYALAQKVSALSALAELEC-----KVALDNALIRAKQAEIDLEAN----- 643
QY 414 RTGFDGLGGRTRIQPDVYMIPLPVKRTKKQKRVKTVSLSQAKID 464
Db 644 -----GRI-----TDLVSN-----NNLTAIKN--KLETELSTAQADLD 675

RESULT 18
MISP_BRUMA STANDARD; PRT; 880 AA.
ID MISP_BRUMA
AC 001202; P90711;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Paramyosin.
OS Brugia malayi.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Brugia.
OX NCBI_TaxID=6279;
RN [1]
RP SEQUENCE FROM N.A.
RA Langy S., Luguidand P., Nicolas L.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RX MEDLINE=92131068; PubMed=175173;
RA Li B., Chandrasekar R., Alvarez R.M., Lifits F., Weil G.J.;
RT "Identification of paramyosin as a potential protective antigen
  against Brugia malayi infection in jirds."
RL Mol. Biochem. Parasitol. 49:315-324(1991).
CC -1- FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF
  MANY THICK FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- SIMILARITY: HIGH, TO MYOSIN HEAVY CHAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL: U77590; AAC18613.1; -.
DR EMBL: M63097; AAA27859.1; -.
DR EMBL: M63098; AAA27860.1; -.
DR HSSP: P80220; IDIP.

```

DR InterPro: IPR002928; Myosin_tail.
 DR Pfam: PF01576; Myosin_tail: 1.
 KW Coiled coil; Muscle protein; Thick filament;
 FT DOMAIN 1 34 NONHELICAL REGION (POTENTIAL).
 FT DOMAIN 35 859 COILED COIL (POTENTIAL).
 FT DOMAIN 860 880 NONHELICAL REGION (POTENTIAL).
 FT CONFLICT 189 189 V -> I (IN REF. 2).
 FT CONFLICT 197 197 L -> A (IN REF. 2).
 FT CONFLICT 251 251 L -> F (IN REF. 2).
 FT CONFLICT 255 255 E -> Q (IN REF. 2).
 FT CONFLICT 696 696 A -> R (IN REF. 2).
 FT CONFLICT 730 736 FRKLRR -> IQEATPA (IN REF. 2).
 FT CONFLICT 825 825 L -> IQ (IN REF. 2).
 FT CONFLICT 857 880 HOLLRKMLOROKETFEKMSNRD -> SSVVTKGNASAK
 FT IYVEDEO (IN REF. 2).
 SQ SEQUENCE 880 AA; 101904 MW; F79A27B642D4A85E CRC64;

Query Match 7.0%; Score 167; DB 1; Length 880;
 Best Local Similarity 19.6%; Pred. No. 0.13;
 Matches 104; Conservative 89; Mismatches 160; Indels 178; Gaps 20;

QY 59 INKLFEDLAOE-----EENVLDREF--LKNELDN---VRAQLSOKDKER-----98
 DB 198 LNRHVNDLAQORRLQAENNNDLKEIHDOKVQDLNQHVKYQLAQOLEEARRLEDAERE 257
 QY 99 -----DSQVIIDPLRDTLEERNATVVSLSQALGKA-----128
 DB 258 RSQQAQOLHOVQVLEDSVFRALADEESAPARAERKLLALANTETIOMKSKFDPAEVALHHEE 317
 QY 129 --EMILCSTLKQMKYLBQOQ---DETROAOEAGRLSKMKT-----166
 DB 318 VEDRRKMKLQKQAEYEDQIEIMLQKISQLEKAKSRLOSEVEVLIVDEKQNTAILERA 377
 QY 167 -EQLELLQSLPREVEEMIRMGVQ-----SAVEQLAVYCVSLAKE 207
 DB 378 KEQLKTVNELKVIIDETLYELAEQAARAALAELOKLMNYEKAVEQKALARENKKL 437
 QY 208 YENLEKARKASGEVADKLK--LDFSSR-----SKLOTYVSEL 243
 DB 438 ODDHEKEKALADANKRLHELDLENARLAGIRELOTLAKSEARDAENKRAQALAEEL 497
 QY 244 DQAKLELKSADKQSDAKDEIMSLKKLTMLQETLNPVAVSETVDR--VLESPAPVE 300
 DB 498 QQLRTM---ERRLOKEEEMELRKNMQF-----EIDRLTALADAEARMK 541
 QY 301 VNKLKRPSPRDDI--DLNATFD-----VDPPARPSSQHGVEKLCLEKSHSPIDVP 353
 DB 542 AEISRLKRRKYQAEIAELMTVDNINRANIEAQTKIKKQSEQLILQASLEDTOROLOOTL 601
 QY 354 KIKCGPKRESQSLGSCSGEPDEELVGAFFIVRNALIGOKPRPRESSESCSDVY 413
 DB 602 DQYALAKRKVSALSALAELEC-----KALDNALIRAKQAEIIDELEAN-----643
 QY 414 RTGFDGLGRTKTIQPTDYVIRPLPVKPKTKVQRQVRVKTPSPSLFOAKID 464
 DB 644 -----GRIT-----TDIVSVN-----NLTALIKN--KLETELSTAQADLD 675

RESULT 19
 MYSX_DROME STANDARD; PRT; 2411 AA.
 AC P05661;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, muscle.
 GN MHC.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89384556; PubMed=2506434;
 RA George E.L., Ober M.B., Emerson C.P. Jr.;
 RT "Functional domains of the Drosophila melanogaster muscle myosin
 heavy-chain gene are encoded by alternatively spliced exons.";
 RL Mol. Cell. Biol. 9:2957-2974(1989).
 RN [2]
 RP SEQUENCE OF 1-312 FROM N.A.
 RX MEDLINE=87280141; PubMed=3038896;
 RA Wassenberg D.R. II, Kronert W.A., O'Donnell P.T., Bernstein S.I.;
 RT "Analysis of the 5' end of the Drosophila muscle myosin heavy chain
 gene. Alternatively spliced transcripts initiate at a single site and
 intron locations are conserved compared to myosin genes of other
 organisms.";
 RT J. Biol. Chem. 262:10741-10747(1987).
 RN [3]
 RP SEQUENCE OF 486-881 FROM N.A.
 RC STRAIN=CANTON-S; TISSUE=Embryonic muscle;
 RX MEDLINE=91330870; PubMed=1907912;
 RA Kronert W.A., Edwards K.A., Roche E.S., Wells L., Bernstein S.I.;
 RT "Muscle-specific accumulation of Drosophila myosin heavy chains: a
 RT splicing mutation in an alternative exon results in an isoform
 RT substitution.";
 RL EMBO J. 10:2479-2488(1991).
 CC - FUNCTION: MUSCLE CONTRACTION.
 CC - SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC - SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC - ALTERNATIVE PRODUCTS: AS DROSOPHILA HAS A SINGLE MUSCLE MHC GENE,
 CC MHC ISOFORMS MUST RELY ON ALTERNATIVE RNA SPLICING THAT LEADS TO
 CC DIFFERENCES IN THE C-TERMINUS OF THE VARIOUS MHC PROTEINS.
 CC - MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC - SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC - SIMILARITY: CONTAINS 1 IO DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M61229; AAA28686.1; ALT_SEQ.
 CC EMBL; M61229; AAA28687.1; ALT_SEQ.
 CC EMBL; J02788; AAA28706.1; ALT_SEQ.
 CC EMBL; J02788; AAA28707.1; ALT_SEQ.
 CC EMBL; X60196; CAA42752.1; ALT_SEQ.
 CC EMBL; X60196; CAA42753.1; ALT_SEQ.
 CC EMBL; X60196; CAA42754.1; ALT_SEQ.
 CC PIR; A28492; A28492.
 CC PIR; A28491; A32491.
 CC PIR; B32491; B32491.
 CC HSP; P08799; LMND.
 CC FlyBase; FBgn0002741; Mhc.
 CC InterPro: IPR000048; IO.
 CC InterPro: IPR004009; Myosin_N.
 CC InterPro: IPR002928; Myosin_tail.
 CC InterPro: IPR001609; myosin_head.
 CC Pfam; PF00612; IO; 2.
 CC Pfam; PF00633; myosin_head; 7.
 CC Pfam; PF02736; Myosin_N; 1.
 CC Pfam; PF01576; Myosin_tail; 1.
 CC PRINTS; PR00193; MYOSINHEAVY.
 CC ProDom; PD0000355; myosin_head; 5.
 CC SMART; SM00015; IO; 1.
 CC SMART; SM00242; MYSC; 1.

DR PROSITE; PS50096; IQ; 1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KM ATP-binding; Calmodulin-binding; Alternative splicing;
 KM Multigene family.
 FT DOMAIN 1 2
 FT DOMAIN 1205 1232
 FT DOMAIN 1233 2376
 FT DOMAIN 227 234
 FT VARSPIC 2385 2385
 FT VARSPIC 2386 2411
 FT CONFLICT 43 44
 SQ SEQUENCE 2411 AA; 276434 MW; 8C4FEC1BB8D310A9 CRC64;

Query Match 7.0%; Score 167; DB 1; Length 2411;
 Best Local Similarity 20.5%; Pred. No. 0.42; Mismatches 144; Indels 140; Gaps 15;
 Matches 93; Conservative 76;

QY 66 LAQEEENVLDREFLNELNDVRAQLSQKDEK-----RDSQVITDITRDITL----- 112
 DB 1999 LEQENNVLAQL---ELSGVROEIDRIQEKKEEFENTKKHQRALDSQASLEAKG 2055
 QY 113 -----ERNATVVSLSQALGKAEMLCSTLKKOMKYLEQOODETKQAQDEAGRLSKM 163
 DB 2056 KAEALRKMKLEADINELTALDHANKANAQAQKIKRYOQOLDIOTALEEQORARDA 2115
 QY 164 KTMQQLL-----LQSOLPEVEEMIRDMGVGSAVEQLAVYVSLKKEENLKEARKA 217
 DB 2116 R--EQGISERRANALONELESRTLEQADRGRQAQ-----ELADAEQJNEVSAQ 2167
 QY 218 SGEVADLRKDLFSSRSKLTQVSELDQ-----AKL--ELKSAQKD 256
 DB 2168 NASISAKRK---LSELDQTHSDDELINLEAKNSEKAKKANVADARLADDELRAQDH 2223
 QY 257 LQSDAKTMSLKKLMLQETLNPV-----ASEYDRLVLE-----SPAP 298
 DB 2224 AQTEKLRKALQEQIKELQVRLDEAEANALGKGAIOKLEQVRELENELEIDGEORRHAD 2283
 QY 299 VEVL-----KLRRSPRFDIDLNATFDVTPPARPSSQHGVEYKLEKLSHSPIQDVP 353
 DB 2284 AQKNLRKSERVVKELSFQSEED-----RKNHRRMODLV 2316
 QY 354 KIKCKGPR-----KESQLSLGQSGAGPEDELVGAFPIFVENAI 393
 DB 2317 DKLOQKIKTKYKQTEAEAEIALMLAKFRKAQQLFEAEERADLAEQAISFRKAGNAGS 2376
 QY 394 LGQKQPKRPSSESCSDVVRTGDCIGGRTKF 426
 DB 2377 VGRGASPAAPRATS-----VRPQFDGLAFPPRF 2403

RESULT 20
 PAM_STRPY STANDARD; PRT; 388 AA.
 AC P49054;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Plasmidogen-binding protein pam precursor (Fragment).
 GN PAM.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 32-41.
 RC STRAIN=AP53 / Serotype M53;
 RX MEDLINE=94064605; Pubmed=8244975;
 RA Berge A., Sjoerding U.;
 RT "PAM, a novel plasmidogen-binding protein from Streptococcus
 pyogenes";
 RU J. Biol. Chem. 268:25417-25424(1993).
 CC -!- FUNCTION: BINDS TO HUMAN PLASMINOGEN. COULD PROVIDE THE BACTERIA

CC WITH A MECHANISM FOR INVASION.
 CC -!- SUBCELLULAR LOCATION: CELL WALL ASSOCIATED.
 CC -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
 CC IN THE REGION OF THE MEMBRANE ANCHOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z22219; CAAB0222.1; -.
 DR HSSP: P03069; IGCL.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR003345; M_repeat.
 DR Pfam: PF02370; M; 8.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
 KM Signal: Repeat; Cell wall.
 FT NON_TER 1 1
 FT SIGNAL <1 19
 FT CHAIN 20 >388
 FT DOMAIN 91 116
 FT REPEAT 91 103
 FT REPEAT 104 116
 FT DOMAIN 147 161
 FT REPEAT 147 153
 FT REPEAT 154 161
 FT DOMAIN 163 278
 FT REPEAT 163 204
 FT REPEAT 205 246
 FT REPEAT 247 278
 FT DOMAIN 344 380
 FT DOMAIN 381 386
 FT NON_TER 388 388
 SQ SEQUENCE 388 AA; 43629 MW; EEECAF9D62CCDB12 CRC64;

Query Match 7.0%; Score 166; DB 1; Length 388;
 Best Local Similarity 20.2%; Pred. No. 0.059;
 Matches 73; Conservative 62; Mismatches 120; Indels 106; Gaps 11;

QY 65 DLQEEENV-----LDREF--LKN-----ELDVNRQAQLSQKDEKRDVSIID 105
 DB 70 DLREKEELQGLKDDVEKLTADAELORLKNERHEAELEKLSERHDHDKKEARKALED 129
 QY 106 TLRTDLEERNATVVSLSQALGKAEMLCSTLKKOMKYLEQOODETKQAQDEAGRLSKMT 165
 DB 130 KLADKQEHNGA-----LRYINKEAEKAEKAE---QKTLKE 164
 QY 166 MEQIELLOSOLPEVEEMIRDMGVGSAVQLAVYVSLKKEENLKEARRASGEVADKL 225
 DB 165 EKQISDASROGLR-----BDLDASREARKQVEKDLNLTAEIDKVKAEKQISDASROGL 218
 QY 226 RKDLFSSRSK-----LQYVSELDQAKLE-----LKSQAKDL 257
 DB 219 RRDIDASREARKQVEKGLANTAEIDKVKAEKQISDASROGLRDLASREARKQVEKAL 278
 QY 258 QSADEKMSLKKLMLQETLNPVASETVDRVLVLESPAVEVNLKLRPSFRDDIDLN 317
 DB 279 EANSKTLAEKLNKLELES-----KLTKEKKEKQAKLEAEKALKLEQJLAKQ 327
 QY 318 A-----TFDVTPPARPSSQHGVEYKLEKLSHSPIQDVPKICKGPKESQLS 367
 DB 328 AEELAKLRAEKASDSQTPPAKPGN-----KAVPGKGAQOAGCTKPNQNKAPKMETKQ 380
 QY 368 L 368
 DB 381 L 381

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U75316; AAB37320.1; -
 DR HSSP: P08799; ILVK.
 DR InterPro: IPR000048; IO.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IO; 2.
 DR Pfam: PF00663; myosin_head; 1.
 DR Pfam: PF02736; myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PSS0096; IO; 1.
 DR Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Multigene family;
 KW Calmodulin-binding.
 FT DOMAIN 1 780 MYOSIN HEAD-LIKE.
 FT DOMAIN 839 1935 COILED COIL (POTENTIAL).
 FT NP_BIND 178 185 ATP (POTENTIAL).
 FT DOMAIN 655 677 ACTIN-BINDING.
 FT DOMAIN 757 771 ACTIN-BINDING.
 FT MOD_RES 129 129 METHYLATION (TRI-) (POTENTIAL).
 FT MOD_RES 695 695 ALKYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 705 705 ALKYLATION (SH-2) (POTENTIAL).
 SQ SEQUENCE 1935 AA; 223110 MW; 149CDBFD910DBB08 CRC64;

Query Match 7.0%; Score 166; DB 1; Length 1935;
 Best Local Similarity 22.2%; Pred. No. 0.37;
 Matches 74; Conservative 62; Mismatches 131; Indels 66; Gaps 9;

QY 79 LKNEIDNVAQLSQKREKRDSDYITLRLDTEE---RNATVVSLOALGAKEMLCST 134
 DB 1287 LSRQLEDEALISQLTGKLTYYTOOLEDLKRLQLEEVYAKNALHQAQSAHNAHDLREQ 1346
 QY 135 LKKQKTYLEQOODETKQAQEEAGRLRSKMT-----MQETILLQSQPREVEEMIRD 186
 DB 1347 YEEETETKAELOAVLSKANSEVAQMRKYEYDIAIQTEELEAKKLAQRLQDAEE--- 1402
 QY 187 MGVOQSAVEQOLAVYCVSLK-----EYENILK---EARKASGEVADKLRDLFSSRSKL 236
 DB 1403 -----AVEAVNAKCSLEKTKHRLQNEIEDLMDVENSMAAALDKKQKNPKIILAEW 1456
 QY 237 QTVVSELDQALFKSAQDLQADKEIMSLKKRL-----TMLQETLNLPPVASETV 288
 DB 1457 KQKTEE---SQSELESSQKEARSLSTELFKLNAYEESLEHLETSKRNKNILQEEISDLT 1513
 QY 289 DRVLESAPPEVNVNKLKRPSPFRDDIDNATFDDVTPPARSSQSHGYEKLCEKSSP 348
 DB 1514 EQLSSSGKTHLEL-KYKQLEAEKLELQS-----ALEESEAS 1550
 QY 349 IQDVPRKICKGPRKESQSLGQSCAGPEDEEL 381
 DB 1551 LEHEEGKILRAQLEFNQIKAEKRLAEKDEEM 1583

RESULT 23
 MYH1_HUMAN STANDARD; PRT; 1939 AA.
 AC P12882; O9Y622;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin heavy chain, skeletal muscle, adult 1 (Myosin heavy chain
 DE IIX/d) (MyHC-IIX/d).
 GN MYH1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Skeletal muscle;
 RX MEDLINE=99318869; PubMed=10388558;
 RA Weiss A., Schiaffino S., Leinwand L.A.;
 RT "Comparative sequence analysis of the complete human sarcomeric myosin
 RT heavy chain family: implications for functional diversity.";
 RL J Mol. Biol. 290:61-75(1999).
 RN [2]
 RP SEQUENCE OF 1064-1939 FROM N.A.
 RX MEDLINE=0616778; PubMed=2421254;
 RA Saez L., Leinwand L.A.;
 RT "Characterization of diverse forms of myosin heavy chain expressed in
 RT adult human skeletal muscle.";
 RL Nucleic Acids Res. 14:2951-2969(1986).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

FT MOD_RBS 709 709 ALKYLATION (SH-2) (POTENTIAL).
FT CONFLICT 1131 1131 A -> T (IN REF. 2).
FT CONFLICT 1139 1139 G -> L (IN REF. 2).
FT CONFLICT 1158 1158 G -> V (IN REF. 2).
FT CONFLICT 1163 1163 A -> T (IN REF. 2).
FT CONFLICT 1286 1289 TESG -> ONOV (IN REF. 2).
FT CONFLICT 1302 1303 VS -> ET (IN REF. 2).
FT CONFLICT 1451 1451 R -> T (IN REF. 2).
FT CONFLICT 1470 1470 E -> V (IN REF. 2).
FT CONFLICT 1473 1474 AS -> SF (IN REF. 2).
FT CONFLICT 1559 1559 L -> V (IN REF. 2).
FT CONFLICT 1598 1598 V -> E (IN REF. 2).
FT CONFLICT 1606 1606 D -> N (IN REF. 2).
FT CONFLICT 1643 1643 A -> D (IN REF. 2).
FT CONFLICT 1648 1648 R -> Q (IN REF. 2).
FT CONFLICT 1750 1750 Q -> K (IN REF. 2).
FT CONFLICT 1822 1822 R -> K (IN REF. 2).
FT CONFLICT 1845 1845 R -> H (IN REF. 2).
SQ SEQUENCE 1939 AA; 223114 MW; 39ADB26AB79DFA53 CRC64;

```

Query Match 7.0%; Score 166; DB 1; Length 1939;

Best Local Similarity 20.8%; Pred. No. 0.37; Indels 88; Gaps 16;

Matches 84; Conservative 83; Mismatches 149; Indels 88; Gaps 16;

```

QY 31 LQCLQSFETAPSRTPQPCRIQVQKRTTINKLFEDLAGEENVLDREF----- 78
DQ 1147 LEESIERLEAGATSAQIEEMNKKEAFQKMRDL--EENATQHEATATLRKKNHDS 1203
QY 79 ---LKNELDNVRAQLSKDKRDSQVILITRLTLERNATVVSLOALGKAEMLCSTL 135
DQ 1204 VAEIGQSDINQIRVQKLEKESEKMEIDLASNME-----TVSKAKGNLEKMCAL 1256
QY 136 KQKMYLEQOODETR-----QAOEAGRL-----BSKKMTMROI 169
DQ 1257 EDQLSEITKEEBOORLINDTAQRARLQTSSEYSRLDEKDTLVLSLQSGKAFTQOI 1316
QY 170 ELLQSQLEPEVEEMIRDMQSGSAVEQQLAVYCVSLKEKYEMLKARKASGEVADKLKRD 229
DQ 1317 EELKR---QLEEEIKAKSALAHALQSSRHDCDLLREQYEEQEKAKELQRAMSKANSEV 1372
QY 230 FSSSKSIOT---VSELDQKKLEKSNOKLOQSDAKRIMSKKLTMLQETLNLPPVAS 285
DQ 1373 AQMTKKTETDAIORTTELEAKKRL--AQR-LQDAEEHVEVANKCASLETKQ--RLQN 1427
QY 286 ETVDRIV-LTSPAPVEVNLKLRPSFRDDI-----DLNATFDVDTPPARPSSO-- 333
DQ 1428 EVEDLMIDVERTNACALADKKQNF-DKILAEMKQCEETHAELEASQKRSRLSTELF 1486
QY 334 ---HGYEKL---CLEKSHSPIQ---DVPKTKCKGPRKESOL 366
DQ 1487 KIKNAYEESLDQLETLKRENKNLQOEISDLTEQIAGEGKRRIHEL 1530

```

```

RESULT 24
MYH9_HUMAN STANDARD; PRT; 1960 AA.
ID AC P35579; O60805;
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Myosin heavy chain, nonmuscle type A (cellular myosin heavy chain,
  type A) (Nonmuscle myosin heavy chain-A) (NMHC-A).
GN MYH9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheta; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewich R., Beare D.M.,
  Clamp M., Smitk L.J., Alnough R., Almeida J.P., Babbage A.,

```

```

RA Baguley C., Bailey J., Barlow K., Bates K.N., Beasley O., Bird C.P.,
  Blakey S., Bridgeman A.M., Buck D., Burgess J., Burrill W.D.,
  Burton J., Carder C., Carter N.P., Chen Y., Clark G., Clegg S.M.,
  Cobley V., Cole C.G., Collier R.E., Connor R.E., Conroy D., Corby N.,
  Coville G.J., Cox A.V., Davis J., Dawson E., Dhanl P.D., Dockree C.,
  Dodsworth S.J., Durbin R.M., Ellington A., Evans K.L., Fey J.M.,
  Fleming K., French L., Garner A.A., Gilbert J.G.R., Goward M.E.,
  Graham D., Griffiths M.N., Hall C., Hall R., Hall-Tamlyn G.,
  Heathcott R.W., Ho S., Holmes S., Hunt S.E., Jones M.C., Kershaw J.,
  Kimberley A., King A., Laird G.K., Langford C.F., Laversha M.A.,
  Lloyd C., Lloyd D.M., Martyn I.D., Mashreghi-Mohammadi M.,
  Matthews L., Mccann O.T., Mcclay J., McLaren S., Mcmurray A.A.,
  Milne S.A., Mortimore B.J., Odell C.N., Pavitt R., Pearce A.V.,
  Pearson D., Phillimore B.J., Phillips S.H., Plumb R.W., Ramsay H.,
  Ramsey Y., Rogers L., Ross M.T., Scott C.E., Sehra H.K., Stuce C.D.,
  Smalley S., Smith M.L., Soderlund C., Spragon L., Stewart C.A.,
  Stuston J.E., Swann R.M., Vaudin M., Wall M., Wallis J.M.,
  Whiteley M.N., Willey D., Williams L., Williams S., Williamson H.,
  Wilmer T.E., Wilming L., Wright C.L., Hubbard T., Bentley D.R.,
  Beck S., Rogers J., Shimizu N., Minoshima S., Kawasaki K., Sasaki T.,
  Asakawa S., Kudoh J., Shintani A., Shibuya K., Yoshizaki Y., Aoki N.,
  Mitsuyama S., Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S.,
  Do A., Do T., Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S.,
  Lai H., Lao H.I., Lewis S., Qi S., Qian Y., Ray L., Ren Q., Shaull S.,
  Nguyen T., Phan S., Qian S., Qian Y., Ray L., Ren Q., Shaull S.,
  Sloan D., Song L., Wang Q., Wang Y., Wang Z., White J., Williamson D.,
  Wu H., Yao Z., Zhan M., Zhang G., Chissoe C., Murray J., Miller N.,
  Mux P., Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H.,
  Bourne S., Cordes M., Du Z., Fulton L., Goela D., Graves T.,
  Harkin J., Hinds K., Kemp K., Latreille P., Layman D., Ozerzky P.,
  Rohlfing T., Scheet P., Walker C., Wamley A., Wohlmann P., Pepin K.,
  Nelson J., Korf I., Bedell J.A., Hillier L., Maris E., Waterston R.,
  Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Salter S.,
  Budarf M.L., Mcdermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
  Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
  Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
  O'Brien K.P., Wilkinson P., Bodenteich A., Hartman R., Hu X.,
  Khan A.S., Lane L., Tlahun Y., Wright H.,
  "The DNA sequence of human chromosome 22."
  Nature 402:489-495(1999).
  [2]
  RP SEQUENCE OF 1-1337 FROM N.A.
  RX MEDLINE=92003925; PubMed=1912569.
  RA Toothaker L.E., Gonzalez D.A., Tung N., Lemons R.S., Le Beau M.M.,
  Arnaout M.A., Clayton L.K., Tenen D.G.;
  "Cellular myosin heavy chain in human leukocytes: isolation of 5'
  cDNA clones, characterization of the protein, chromosomal
  localization, and upregulation during myeloid differentiation."
  Blood 78:1826-1833(1991).
  [3]
  RP SEQUENCE OF 1-715 FROM N.A.
  RX MEDLINE=91316803; PubMed=1860190.
  RA Simons M., Wang M., McBride O.W., Kawamoto S., Yamakawa K.,
  Gdula D., Adelsstein R.S., Weir L.;
  "Human nonmuscle myosin heavy chains are encoded by two genes located
  on different chromosomes."
  J. Clin. Res. 69:530-539(1991).
  [4]
  RP SEQUENCE OF 714-1960 FROM N.A.
  RX MEDLINE=90138958; PubMed=1967835.
  RA Saez C.G., Myers J.C., Shows T.B., Leitwand L.A.;
  "Human nonmuscle myosin heavy chain mRNA: generation of diversity
  through alternative polyadenylation."
  Proc. Natl. Acad. Sci. U.S.A. 87:1164-1168(1990).
  [5]
  RP VARIANT DFNAL7 HIS-705.
  RX MEDLINE=2049856; PubMed=11023810;
  RA Ialwani A.K., Goldstein J.A., Kelley M.J., Luxford W., Castelein C.M.,
  Mhatre A.N.;
  "Human nonsyndromic hereditary deafness DFNAL7 is due to a mutation in
  nonmuscle myosin MYH9."
  Am. J. Hum. Genet. 67:1121-1128(2000).
  [6]

```

RP VARIANTS MHA/FTNS/SBS K-93; C-702; C-1165; H-1424 AND K-1841.
RX MEDLINE-20428192; PubMed-10973259;
RA Seri M., Cusano M., Gangarosa S., Caridi G., Bordo D., Lo Nigro C.,
RA Ghigeri G.M., Ravazzolo R., Savino M., Del Vecchio M., d'Apolito M.,
RA Iolascon A., Zelante L.L., Savoia A., Baldini C.L., Norris P.,
RA Magrini U., Bellelli S., Heath K.E., Babcock M., Glucksman M.J.,
RA Aliprandis E., Bizzaro N., Desnick R.J., Martignetti J.A.,
RT "Mutations in MYH9 result in the May-Hegglin anomaly, and Fechtner and
RT Sebastian syndromes";
RL Nat. Genet. 26:103-105(2000).
RN [7]
RP VARIANTS MHA TLE-1155 AND LYS-1841.
RX MEDLINE-20428193; PubMed-10973260;
RA Kelley M.J., Jawien W., Ortel T.L., Korczak J.F.;
RT "Mutation of MYH9, encoding non-muscle myosin heavy chain A, in
RT May-Hegglin anomaly";
RL Nat. Genet. 26:106-108(2000).
CC -I- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
CC CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
CC CAPPING.
CC -I- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
CC CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
CC REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -I- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -I- DISASE: DEFECTS IN MYH9 ARE THE CAUSE OF MAY-HEGGLIN ANOMALY
CC (MHA), AN AUTOSOMAL DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED
CC BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUKOCYTE INCLUSIONS.
CC -I- DISASE: DEFECTS IN MYH9 ARE THE CAUSE OF FECHTNER SYNDROME
CC (FTNS), AN AUTOSOMAL DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED
CC BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUKOCYTE INCLUSIONS.
CC WITH ADDITIONAL ALPORT-LIKE CLINICAL FEATURES OF SENSORINEURAL
CC DEAFNESS, CATARACTS AND NEPHRITIS.
CC -I- DISASE: DEFECTS IN MYH9 ARE THE CAUSE OF SEBASTIAN SYNDROME
CC (SBS), AN AUTOSOMAL DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED
CC BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUKOCYTE INCLUSIONS.
CC -I- DISASE: DEFECTS IN MYH9 ARE THE CAUSE OF AN AUTOSOMAL DOMINANT
CC FORM OF NONSYNDROMIC SENSORINEURAL DEAFNESS (DNA17) WHICH IS
CC CHARACTERIZED BY PROGRESSIVE HEARING IMPAIRMENT AND
CC COCHLEOSACULAR DEGENERATION.
CC -I- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: 282215; CAB05105.1; -
DR EMBL: M01105; AAA59888.1; -
DR EMBL: M69180; AAA61765.1; -
DR EMBL: M31013; AAA6349.1; -
DR HSSP: P08799; 1LVK.
DR MIM: 160775; -
DR MIM: 153640; -
DR MIM: 155100; -
DR MIM: 603622; -
DR MIM: 605249; -
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR011609; myosin_head.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF02736; myosin_head; 1.
DR Pfam: PF02236; myosin_N; 1.
DR Pfam: PF01576; myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.

DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
KW Coiled coil; Alkylation; Multigene family; Disease mutation;
KW Deafness.
FT DOMAIN 1 778
FT DOMAIN 779 808
FT DOMAIN 837 1926
FT NP_BIND 174 181
FT DOMAIN 654 676
FT MOD_RES 694 694
FT MOD_RES 704 704
FT VARIANT 93 93
FT VARIANT 702 702
FT VARIANT 705 705
FT VARIANT 1155 1155
FT VARIANT 1165 1165
FT VARIANT 1424 1424
FT VARIANT 1841 1841
FT CONFLICT 53 55
FT CONFLICT 660 660
FT CONFLICT 869 869
FT CONFLICT 931 931
FT CONFLICT 1240 1241
FT CONFLICT 1350 1350
FT CONFLICT 1764 1764
FT CONFLICT 1771 1771
SQ SEQUENCE 1960 AA; 226531 MW; 588BF4BB8C106E6F CRC64;
Query Match 7.0%; Score 166; DB 1; Length 1960;
Best Local Similarity 19.6%; Pred. No. 0.37;
Matches 88; Conservative 81; Mismatches 145; Indels 136; Gaps 15;
QY 48 OCRIVGKRTIINKLFEDFAOEENVIDREF-----LKNELDNVRAQLSOK 93
DB 1073 ELKMQLAKEE--BELQALALARVEEAAQNMALKRTRELSQSLQEDLESRSARRNA 1130
QY 94 DKERDSQVYITDRLPTLEBRNATVVSLODALGAE---MLCSTLKKQMKYIEQOQDET 149
DB 1131 EKQKRDIGETELAKTELEDYTDSTAQOELRSKREQEVNIIKKTLEERAKTHEAQIOEM 1190
QY 150 KQKQEEA-----GRLSKMKTYE-----QTELLLOS-----QLPEY 180
DB 1191 RQKHSQAVELEAQLDQTKRKVKANLEKAKOTLENERGELANEKVALDQKGDSERHRRKV 1250
QY 181 EEMIRDMGV-----GQSAVEQLAVYCVSLKKEYENLKEARKASGEVADLKRLKDFSSRSKL 236
DB 1251 EAQLQELQYKFNENGERVRELDKVKTKLOVELDNVYGLLSQSSKSKLTKQSALESQSL 1310
QY 237 QIVYSELDAQKLELKSQKQDQASADKEIMSLKKKLTMLQETLNLPPVASETYDRLVLESP 296
DB 1311 Q-----PTQELQELQENRQKLSLSTYKQVEDEKN----- 1339
QY 297 APVEVNLKLRPSFRDIDINATFDVTPPARSSSOHGVEKLECLKS-----HSPIDQY 352
DB 1340 -----SFRQLE-----EEEEKHN-----LEKQIATLHAQVADM 1369
QY 353 PKRI-----CKGPRKESQSLSGQSCAGCEPDELVGAFFIFRNAILQCKQ----- 398
DB 1370 KKKMEDSVGLTFAEYKRRKQLQKDLGLS---QRHEKVAAYKLTETRTRLQQLDLDLL 1426
QY 399 -PKRPSSESSCKDQVVRVTFGDLGGRTKFI 427
DB 1427 VLDLHQRSACNLEKKQKFKDQLAEKTI 1456

FT VARSPLIC 2441 2704 MISSING (IN ISOFORM 4).
 FT CONFLICT 1998 1998 G -> R (IN REF. 5).
 SO SEQUENCE 2704 AA: 313084 MW: A7219E687A634A77 CAC64:

Query Match 7.0%; Score 166; DB 1; Length 2704;
 Best Local Similarity 20.0%; Pred. No. 0.54;
 Matches 97; Conservative 88; Mismatches 181; Indels 120; Gaps 19;

QY 33 CLDSEFTASRCPQRIQVGRITIKLFPDLQAEENVLDREFLNKELNDVRA----- 88
 DB 1385 CRENAFLPCVITQATSGRAVATG-----LQEHKQKAEELKQVDELTAANRK 1432
 QY 89 -----QLSQDKERDSQVITDRLDLEERNATVVSLOQALGKAEMLSTLKKQ 138
 DB 1433 AEDQMRELTYELNALQLEKSSSEKALLDKLDETNTNTRLCLELEKRDQAEKGSQO 1492
 QY 139 MKYLEQOQDET---KQAEAGRLRSKMTMEQIEL-----LLQSQLPEVEEMIRD 186
 DB 1493 LRELGRQINOTGTGAERAMQDASDLK-KIKRNYQLELESINHEKGLQR---EVDRTIRA 1548
 QY 187 MGVOGSANEVDL-AVYCVSLKKEVENKEARKASGEVADLRDLFSSRSKLVYSELD 244
 DB 1549 HAAVEKNIOHLNQSIOHSFRDEKELERLQIQRSMDLKEQFEK---SHEOLLQNIKAKE 1605
 QY 245 QAKLELSAQKDLSADKEIMSLKKLTMLOETLNLPPVASETV---DRLVLESPARVE 300
 DB 1606 NND-KIQRNLMELEKSENECKAEMLKQKVEELTRONNETKIMQRIQASSENIVLEKQITIQ 1664
 QY 301 -VNLIKRRSFRRDIDLNAFTVDTPPARPSSQHGVEYKLEKLSHSPIDVPPKICK 358
 DB 1665 RCEALKTQADGFKDQL-----RSTNEH-----LHKOTKTBDQFORKI-- 1701
 QY 359 GPRKESLSLGGGSCAE---PDELYGAP-----IVRNAIILQOKPQRSSSSC 408
 DB 1702 -----KCLEEDLAKSQNLVSEFKQKCDQOIIITQNT---KKEVRNLMAELNA 1745
 QY 409 SKDVVATGFDGLGRTKTIQPTDTVMIRPLPVKPKTKQKORVRYKTVPS-----LFO 460
 DB 1746 SKEKRRGEGQKVOLOQAOVELNRL-----KKVQDELHKTIEQMTIHKMYLFO 1796
 QY 461 AKLDTE 466
 DB 1797 EESGKF 1802

RESULT 26
 USOL_YEAST STANDARD; PRT; 1790 AA.
 ID USOL_YEAST
 AC P25386;
 DT 01-MAY-1992 (Rel. 22, Last Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Intracellular protein transport protein USOL.
 GN USOL OR INT1 OR YDL058W.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID:4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180-1A;
 RX MEDLINE=91185402; PubMed=2010462;
 RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
 RA Yamasaki M.;
 RA "A cytoskeleton-related gene, usol, is required for intracellular
 RT protein transport in Saccharomyces cerevisiae.";
 RL J. Cell Biol. 113:245-260(1991).
 RN [2]
 RP SEQUENCE OF 782-1790 FROM N.A.
 RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
 RA Kendrick K.E.;
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE OF 1-8 FROM N.A.
 RA Bai Y., Symington L.S.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
 CC ER AND THE GOLGI COMPLEX.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
 CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
 CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
 CC -1- SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X54378; CA38253.1; -;
 CC EMBL: L03188; AAB00143.1; -;
 CC EMBL: U53668; AAB66659.1; -;
 CC PIR: A38455; A38455.
 CC HSP: P80220; 1DIP.
 CC SGD: S0002216; USOL.
 CC InterPro: IPR002017; Spectrin.
 CC TransPro: Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
 CC
 CC FT DOMAIN 1 724 GLOBULAR HEAD.
 CC FT DOMAIN 725 1790 COILED COIL (POTENTIAL).
 CC FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).
 CC FT DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.
 CC FT DOMAIN 1172 1786 ASP/GLU-RICH (ACIDIC).
 CC FT CONFLICT 847 847 G -> E (IN REF. 2).
 CC FT CONFLICT 924 924 E -> K (IN REF. 2).
 CC FT CONFLICT 1253 1253 V -> I (IN REF. 2).
 CC FT CONFLICT 1319 1319 I -> V (IN REF. 2).
 CC FT CONFLICT 1461 1461 N -> S (IN REF. 2).
 CC FT CONFLICT 1581 1581 G -> S (IN REF. 2).
 CC FT CONFLICT 1600 1600 I -> V (IN REF. 2).
 CC FT CONFLICT 1661 1661 R -> S (IN REF. 2).
 CC FT CONFLICT 1772 1772 D -> DEEDDEE (IN REF. 2).
 CC FT SEQUENCE 1790 AA: 206424 MW: 6CE2B216E9PD4818 CAC64:

Query Match 6.9%; Score 165.5; DB 1; Length 1790;
 Best Local Similarity 20.9%; Pred. No. 0.36;
 Matches 78; Conservative 75; Mismatches 129; Indels 91; Gaps 13;

QY 55 KRTINKLFPDLQAO---EEENVLDREFLNK-----NEIDNVRAQLSQDKKKRDSQV 102
 DB 1368 ERKLLNGSSPTIQEYSEKWTLEDELIRLQENNELKAKEIDNTRSEL-EKVSLSND--- 1423
 QY 103 IIDRLPTLEERNATVVSLOQAL-----GKAEMLCSTLKKMKYLEQOQDETQAOE 154
 DB 1424 -----ELLEKQNTISLQDEITLSYDKDKTRNDEKLSLTERDNKRDESLKEQLRAQE 1477
 QY 155 EAGRLRSKMTMEQIELLSQLPVEEMIRDMVGOSAVE---QLAVYCVSLKKEVEN 210
 DB 1478 SKAVEGGLKLEESSSEKAEKELSKEMMKL---ESTIESNTEELKSSMETIRKDEK 1534
 QY 211 LKEARKASGEVADKLK-----KDLFSSSKLQTVY---SELDQAKLEIKSA 253
 DB 1535 LEQSKSAEEDIKMLQHEKSDLSIRINESEKDIELKSLRIEAKSGSELETVQOELNNA 1594
 QY 254 QKDQASDKETMSLKKLTMLOETLNLPPVASEVDRVLESPPAVPN-----LKLRR 307
 DB 1595 QEKTRINAEENTVYLSKL-----EDIRELKDQAEIRISNOEKEKLLYSRL 1640
 QY 308 PSFRDDIDLNAFTVDTPPARPSSQHGVEYK-----LCLEKSHSPI-----Q 350

Db 1641 KEBLELDSTOOKAKOSEERRAEVRFQVEKSOLEKAMLETETKYNDLVNKQAMKDE 1700
QY 351 DVPKICKGPRKE 363
Db 1701 DTVAKTTDSOROE 1713

RESULT 27
MYSR_DIRIM
ID_MYSR_DIRIM STANDARD: PRT: 848 AA.
AC P13392;
DT 01-JUN-1990 (Rel. 13, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Paramyosin (Fragment).
OS Dirofilaria immitis (Canine heartworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Dirofilaria.
OX NCBI_TaxID=6287;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90220759; PubMed=2325708;
RA Limberger R.J., McReynolds L.A.;
RT "Filarial paramyosin: CDNA sequences from Dirofilaria immitis and
RT Onchocerca volvulus."
RL Mol. Biochem. Parasitol. 38:271-280(1990).
RN [2]
RP SEQUENCE OF 18-96 FROM N.A.
RX MEDLINE=89344126; PubMed=2527335;
RA Granda A.G. III, Tuyen L.K., Asiklin N., Davis T.B., Philipp M.,
RA Cohen C., McReynolds L.A.;
RT "A lambda gt11 cDNA recombinant that encodes Dirofilaria immitis
RT paramyosin."
RL Mol. Biochem. Parasitol. 35:31-41(1989).
CC -1- FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF
CC MANY THICK FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- SIMILARITY: HIGH, TO MYOSIN HEAVY CHAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL: M29733; AAA28299.1; -;
DR EMBL: J04009; AAA28300.1; -;
DR PIR: A45548; A45548;
DR PIR: A44972; A44972;
DR HSSP: P80220; 1DIP;
DR InterPro: IPR002928; Myosin_tail.
DR Pfam: PFO1576; Myosin_tail; 1.
KW Coiled coil; Muscle protein; Thick filament; Myosin.
FT NON_TER 1
FT DOMAIN <1 9 NONHELICAL REGION (POTENTIAL).
FT DOMAIN 10 833 COILED COIL (POTENTIAL).
FT DOMAIN 834 848
FT DISULFID 105 105 NONHELICAL REGION (POTENTIAL).
FT DISULFID 594 594 INTERCHAIN (POTENTIAL).
FT NON_TER 848 848 INTERCHAIN (POTENTIAL).
SQ SEQUENCE 848 AA; 98008 MW; BF32B7120157FA50 CRC64;

Query Match 6.9%; Score 165; DB 1; Length 848;
Best Local Similarity 19.2%; Pred. No. 0.16;
Matches 102; Conservative 90; Mismatches 161; Indels 178; Gaps 19;

QY 59 INKLFPLAD-----EENVLDREF--LNKELDN---VRAQLSQKKEK----- 98
Db 171 LNRHVNDLAQORORLQAEENNDLKEIHDKVYQVOLDNIQHVYQVLAQOLEFAARRLEDAERE 230

QY 99 -----DSQVITDRLDRTLEERNATVVSLOQALGA----- 128
Db 231 RSQLOAQLOHVOLELDSVRYTALDEESARAEAEKRLALATETIYQWKSKEPDAEVALHHEE 290
QY 129 --EMICSTLKKQMYLYEQOQ---DETQAOEAEAGRLSKMKT----- 166
Db 291 VEDLRKKMLQKQAEYEQIEIMLOKISOLEKAKSRLQSEVEVLIVDEKAQNTAILERA 350
QY 167 -EQIELLOSQLPEVEEIRDMGVQ-----SAVEQLAVYCVSLKE 207
Db 351 KEOLEKTVNELKVRIDELTYVELEAQRARAALAELOKMKNLKALOEKALARENKKL 410
QY 208 YENLEKARASGEVADRLK-DLFSSR-----SKLOTVSEL 243
Db 411 QDDLHEAKEALADANKRLHELDLENARLAGIRELOALAKESEARDAERARALAE 470
QY 244 DQAKLELSAQKDSADKEIMSLKKRLTMLQETLNPVASETVDR/-VLESPAPVE 300
Db 471 QQLRTM---ERRLOKEEEMEARLKNMQF-----EIDRLTALADAEARMK 514
QY 301 VNLKRRPSFRDQI-DLNATFD-----VDTPPARPSSOHGYEKECLEKSHSPIDVP 353
Db 515 AEISRLKKKKQAEIJAELMTVDNLRANIEAQTKIKQSEQLKILQASLEDTORLOQOTL 574
QY 354 KKICKGPRKESQLSLGGOSGAGEPDELVGAFPIFVRNALIGOKQKRRPSESSCSKDVY 413
Db 575 DQYLAQRKYSALSALEEC-----KVALDNALAKARQAEIDLEANA----- 617
QY 414 RTGFDGLGRTKFTPTDVTYMRPLPVKPKTKVQRVRYKTPSLFQAKID 464
Db 618 -----RITDLVLSIN-----NNTLAIRN--KLETELSTAQADLD 648

RESULT 28
RAD5_ARCFU
ID_RAD5_ARCFU STANDARD: PRT: 886 AA.
AC Q29230;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE DNA double-strand break repair rad50 Arpase.
GN RAD50 OR AFI032.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
CC Archaeoglobus.
CC NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kierlavage A.R., Graham D.E., Kypides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlisch P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
CC rad50/mre11 complex possesses single-strand endonuclease activity
CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mre11 by unwinding
CC and/or repositioning DNA ends into the mre11 active site (By
CC similarity).
CC -1- SUBUNIT: Forms a complex with mre11 (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.


```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001032; AAB90211.1; -
DR TIGR: AF1032; -
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR001238; RecF.
DR InterPro: IPR002017; Spectrin.
DR Pfam: PF00470; RecF; 1.
KW DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
FT NP_BIND 31 38 ATP (BY SIMILARITY).
FT DOMAIN 148 728 COILED COIL (POTENTIAL).
SQ SEQUENCE 886 AA; 103653 MW; D35641D499AAB58 CRC64;

Query Match 6.9%; Score 165; DB 1; Length 886;
Best Local Similarity 21.1%; Pred. No. 0.17; Indels 82; Gaps 18;
Matches 97; Conservative 104; Mismatches 176;

QY 59 INKLFPLDAGEENVLDL---EFLKNELDNVRAO-----LSQDKKE-KRDSQVI 103
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 206 IESIREKLESEVRNLESRLKLELNKSLKQESSVLQEVNGLEKLELEKOLKEV 265
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 104 IDTLRDLTEERNATVSLQ---QALGAEMLCSTLKKOMLYLEQO-----ODET 149
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 266 VERIED--LEKKAKVEKLEKAKERYSTLEKLEINQALDVERKEGDLTREAGIQAL 324
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 150 KQAGEAGRLRSKMKKMEQIE-----LILQSLPEVEEMIRMGVQSGASVEQLAV 199
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 325 KKAEDSKSLKETTKRLEERELEREKSHRLLETLPKMDRY---QGI-KAKLEKNL 380
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 200 YCVSLKREYENLEKARRASGEVADLRKDLFSSRSKLQTVYSELQAKLELSAQKDLQS 259
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 381 TPDKVEKMYDLSKAKKEKEKETEKLKK-LIAKKSLSLKTGAOLKKAVEELKSAERCPV 439
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 260 ADKRTMSLKKKLTMLQETLNPVASE---TVDRLVLESAPAVEYNLKRPP--SFRD 312
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 440 CGRLDEHGRNMAEYTRREKRIAEELAKADETEKLEKLEKALEKQETVLKRYQ 499
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 313 DID-LNA-TFDVDPAPRPPSSOHGYEKLCLEKSHSPIDQVPKICKGPKKESQLSLG 370
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 500 MWDELKALENLSHDAEKLSAESEYRKY-----KERLDGLGQOKILLSS 546
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 371 QSCAGEPDEELVGAFLFVRNALIGQKQPKRPSSESSCKDVRTGFDLGGRTKFTOPT 430
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 547 ASRIKEIKSL-----REIEALKNVESEGE--LHKRIREGFSELEELREVQSL 596
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 431 DTVAIRPLPKYP-----KTKYKQVRKYKTVPSLFOAKLD 464
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 597 RPFYNNKMLELKDAESRLESELKREKLEDEISEAIKLE 635
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 29
RA50_PYRAB STANDARD; PRT; 880 AA.
AC Q9U2C8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE DNA double-strand break repair rad50 ATPase.
GN RAD50 OR PAB0812.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RC SEQUENCE FROM N.A. STRAIN=ORSAT;
```

```
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RT structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
CC rad50/mre11 complex possesses single-strand endonuclease activity
CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mre11 by unwinding
CC and/or repositioning DNA ends into the mre11 active site (By
CC similarity).
CC -!- SUBUNIT: Forms a complex with mre11 (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ248286; CAB50131.1; -
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR001687; ATP_GTP_A.
DR InterPro: IPR001238; RecF.
DR Pfam: PF00470; RecF; 1.
DR SMART: SM00382; AAA; 1.
KW DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
FT NP_BIND 30 37 ATP (BY SIMILARITY).
FT DOMAIN 144 745 COILED COIL (POTENTIAL).
SQ SEQUENCE 880 AA; 103970 MW; FDB177EC7E026479 CRC64;

Query Match 6.9%; Score 164.5; DB 1; Length 880;
Best Local Similarity 21.9%; Pred. No. 0.18;
Matches 61; Conservative 65; Mismatches 122; Indels 31; Gaps 7;

QY 65 DLAQEEENVLDREFLKNELDNVRAOQSQDKERKDSQVITLRLDTLEERNATVSLQQA 124
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 226 ELEKVENKVELESIKKISIELKIQVEKLGKRGKGLDEKTVQJERSIEEKAKISELEI 285
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 125 LGKAEMLCSTLKKOM---VLEQODETKQAGEAGRLRSKMKTMEQIEILLQSLPEVE 181
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 286 VKDIPKIQEKEKEVKRLKGRDEYESLRLRLEKLSWESELAIEIVINEGKKKRAE 345
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 EMTRDAGVGASVQQLAVYCSL-----KKEVENLKARK--ASGEVADKL-----RK 227
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 346 EIREKLEIEIKRLBELPYYEELEDARQVOQIERLAKRLGLSPGEVIEKLESEKERT 405
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 DLFSSRSKLQTVYSELQAKLELSAQKDLQSA-----DKETMSLKKKLTMLQETLNL 280
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 406 EIEPAIKEITTRIGQMOQEKNERKKAIEELKKAQKCPVCGRELTTEHKKLEMERITLEI 465
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 281 PVAASETVDRLVLESAPAVEYNL-----KLRRPSFRDII 314
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 466 KRIEEL--KRTTEERKRLRVNLKLEIKLREPSVMNDI 502
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 30
M6_STRPY STANDARD; PRT; 483 AA.
AC P08089;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE M protein, serotype 6 precursor.
GN EMM6.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OX NCBI_TaxID=1314;
RN [1]
```


DR SMART; SM00180; EGF_Lam; 10.
DR SMART; SM00001; EGF_Like; 1.
DR SMART; SM00281; Lamb; 1.
DR SMART; SM00136; LamNT; 1.
DR PROSITE; PS00022; EGF_L; 8.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01248; LAMININ_Type_EGF; 11.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT SIGNAL 1 33 POTENTIAL.
FT CHAIN 34 1639 LAMININ GAMMA-1 CHAIN.
FT DOMAIN 34 298 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 299 358 LAMININ EGF-LIKE 1.
FT DOMAIN 359 413 LAMININ EGF-LIKE 2.
FT DOMAIN 414 460 LAMININ EGF-LIKE 3.
FT DOMAIN 461 513 LAMININ EGF-LIKE 4.
FT DOMAIN 514 523 LAMININ EGF-LIKE 5 (N-TERMINAL).
FT DOMAIN 524 709 LAMININ DOMAIN IV.
FT DOMAIN 710 743 LAMININ EGF-LIKE 6.
FT DOMAIN 744 792 LAMININ EGF-LIKE 7 (C-TERMINAL).
FT DOMAIN 793 846 LAMININ EGF-LIKE 8.
FT DOMAIN 847 901 LAMININ EGF-LIKE 9.
FT DOMAIN 902 955 LAMININ EGF-LIKE 10.
FT DOMAIN 956 1003 LAMININ EGF-LIKE 11.
FT DOMAIN 1004 1049 LAMININ EGF-LIKE 10.
FT DOMAIN 1050 1609 LAMININ EGF-LIKE 11.
FT DOMAIN 1087 1109 COILED COIL (POTENTIAL).
FT DOMAIN 1144 1247 COILED COIL (POTENTIAL).
FT DOMAIN 1306 1627 COILED COIL (POTENTIAL).
FT DISULFID 299 308 BY SIMILARITY.
FT DISULFID 301 322 BY SIMILARITY.
FT DISULFID 324 333 BY SIMILARITY.
FT DISULFID 336 356 BY SIMILARITY.
FT DISULFID 359 368 BY SIMILARITY.
FT DISULFID 361 384 BY SIMILARITY.
FT DISULFID 387 396 BY SIMILARITY.
FT DISULFID 399 411 BY SIMILARITY.
FT DISULFID 414 426 BY SIMILARITY.
FT DISULFID 416 432 BY SIMILARITY.
FT DISULFID 434 443 BY SIMILARITY.
FT DISULFID 446 458 BY SIMILARITY.
FT DISULFID 461 475 BY SIMILARITY.
FT DISULFID 463 482 BY SIMILARITY.
FT DISULFID 484 493 BY SIMILARITY.
FT DISULFID 496 511 BY SIMILARITY.
FT DISULFID 744 753 BY SIMILARITY.
FT DISULFID 746 760 BY SIMILARITY.
FT DISULFID 762 771 BY SIMILARITY.
FT DISULFID 774 790 BY SIMILARITY.
FT DISULFID 793 801 BY SIMILARITY.
FT DISULFID 795 811 BY SIMILARITY.
FT DISULFID 814 823 BY SIMILARITY.
FT DISULFID 826 844 BY SIMILARITY.
FT DISULFID 847 861 BY SIMILARITY.
FT DISULFID 849 868 BY SIMILARITY.
FT DISULFID 871 880 BY SIMILARITY.
FT DISULFID 883 899 BY SIMILARITY.
FT DISULFID 902 919 BY SIMILARITY.
FT DISULFID 926 936 BY SIMILARITY.
FT DISULFID 928 937 BY SIMILARITY.
FT DISULFID 940 953 BY SIMILARITY.
FT DISULFID 956 968 BY SIMILARITY.
FT DISULFID 958 975 BY SIMILARITY.
FT DISULFID 977 986 BY SIMILARITY.
FT DISULFID 989 1001 BY SIMILARITY.
FT DISULFID 1004 1016 BY SIMILARITY.
FT DISULFID 1006 1022 BY SIMILARITY.
FT DISULFID 1024 1033 BY SIMILARITY.
FT DISULFID 1036 1047 BY SIMILARITY.
FT DISULFID 1050 1050 BY SIMILARITY.
FT DISULFID 1053 1053 INTERCHAIN (PROBABLE).
FT DISULFID 1631 1631 INTERCHAIN (PROBABLE).
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 862 862 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 965 965 N-LINKED (GLCNAC. . .) (POTENTIAL).
Query Match 6.9%; Score 163.5; DB 1; Length 1639;
Best Local Similarity 21.7%; Pred. No. 0.4;
Matches 117; Conservative 72; Mismatches 220; Indels 129; Gaps 21;
QY 7 CTISDFHDSRDVAALHCHGTFHLOCLIOSFETAPSRTPQCRIDYQKRTIINKLEFPL 66
DB 1047 CIDPDDCYNLVQDAADLHRAKLEFNLSTQLEIARTPTVNDDEE--AKLAAVEKVAVL 1103
QY 67 AOEEN-----VLDRFLKNELDNVRAQLSOKREKRPDQYIITLRTLEERN 115
DB 1104 AQDARDSDGCGGYAAVEYID--LHKHLDSYREHLVSADKFOADANEIDRAR-----QN 1156
QY 116 ATVV-----SLQALG---KAEMLCSTLKKOMKYLEQOODETKQAQAEAGRLRS 161
DB 1157 YTIIDQITENAKKLELQALDLNDEGAQALARAKEYSVEF--GOOSEQISDISREARALD 1215
QY 162 KMKTMEDTEL-----LLQSPLPEVEEMIRDMGVSQAVRQLAYVCV 202
DB 1216 KLESEAFDLKNAKADAKAVEKAHQAKSAIDQLKIGTFLRESEVGLISHVKQ----- 1269
QY 203 SLKKEYENLKEARKASGEVADKLRLKDLFSSRSKLTQYVSELDQAKLEKSAOKDLOSADK 262
DB 1270 SLGTVQTSKFEALRKANEVVD-----TALTILNDV--NQTQPEIDISOLKDAVANE 1321
QY 263 EIMSLKRLKLTWLOETLMLPVASETVDRLVLESPAPYENLKLRRSPFRDIDL---NA 318
DB 1322 RADELTKQITELNSNGELPADFETQEL-----TEALLKRAEQOQLDIELLERAKA 1374
QY 319 TFDVDTPPARPSSS-----QNGYYEKL-----LEKSHSPIQDVPKRIKCGPRRESQ 365
DB 1375 AHDKATRAVEQDGTLEANNTYEKLAFQSDVORSSSESAKALQTVP-----NIEKEIQ 1429
QY 366 LSLGQSCAGEPDEIVGAPPIEFVRNALIGOKP-----KRPRESS 407
DB 1430 ---NAESLIQAEALDGA---NKNANEAKKNAQEAQLKAYADASDAELIRKANETK 1482
QY 408 CSKDVRTGFDGLGRRTKFTQPTDVTVMIRPLPVKPKTKYKQVRVKTVPSPGAKLDT 465
DB 1483 VAARNLREADDQLNHRVKLTE-MDIFKLEESSTFDMLVYDAKR-----KVGAQAKDT 1534
RESULT 33
ALM1_SCHPO STANDARD; PRT; 1727 AA.
ID ALM1_SCHPO
AC G9UTR5; 013313; G9UTR8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Abnormal long morphology protein 1 (Sp8).
GN ALM1 OR SPAC1486.04C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=972;
RA McDougall R.C., Rajandream M.A., Barrell B.G., Seeger K., Harris D.;
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
RX MEDLINE=20123449; PubMed=10660053;
RC STRAIN=972;
RX MEDLINE=20123449; PubMed=10660053;
RA Jimenez M., Petit T., Gancedo C., Goday C.;
RT "The alm1+ gene from Schizosaccharomyces pombe encodes a coiled-coil
protein that associates with the medial region during mitosis.";
Mol. Gen. Genet. 262:921-930(2000).
N-LINKED (GLCNAC. . .) (POTENTIAL).

```

RN [3]
RP SEQUENCE OF 644-834 FROM N.A.
RC STRAIN-968 H90;
RA Ding D.;
RT "Generation and analysis of GFP-gene fusion library of fission yeast."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: AFTER THE ONSET OF MITOSIS, AT MID- TO LATE ANAPHASE, CO-LOCALIZES WITH THE MEDIAL ACTIN RING. MAY PLAY A ROLE IN CYTOKINESIS.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL13357; CAB62414.1; -
DR EMBL: AF010473; AAB65416.1; ALT_INIT.
DR EMBL: AB028012; BAA87316.1; -.
KW Coiled coil.
FT DOMAIN 57 361 COILED COIL (POTENTIAL).
FT DOMAIN 443 463 COILED COIL (POTENTIAL).
FT DOMAIN 542 740 COILED COIL (POTENTIAL).
FT DOMAIN 804 1106 COILED COIL (POTENTIAL).
FT DOMAIN 1223 1427 COILED COIL (POTENTIAL).
FT DOMAIN 1497 1555 COILED COIL (POTENTIAL).
FT DOMAIN 1601 1664 COILED COIL (POTENTIAL).
SQ SEQUENCE 1727 AA; 197858 MW; F820BF8D9C132644 CRC64;

```

```

Query Match 6.9%; Score 163.5; DB 1; Length 1727;
Best Local Similarity 22.8%; Pred. No. 0.43;
Matches 97; Conservative 60; Mismatches 132; Indels 137; Gaps 17;

```

```

OY 30 HLOCLISFETAPSPRCPCORIO---VGKRTIINKL---PFDLAQEBENV-----LDRE 77
DB 1301 HNOEITLILSNSSITSSDSARLKNELYSKEMLELNOEIGHLKSLEYTKXSSEDLENE 1360
OY 78 FLKN-----ELDN-----VRAOLSQDKERKDSQVIIDT 106
DB 1361 RAQNOQKIEDELKNTKLAAMRTKYEQVYNKSLKHNQIRQOLSQKTSLEAKVACHQ 1420
OY 107 LROTLERNAT-----VVSLOALGKAEMLCSTLKKOMKYL----- 142
DB 1421 LNEOLNKPSPATPTATTOSEPSTVSLSEFNSTKELSTORLSTMDILNTTKEELEKVR 1480
OY 143 -----EODODETKOAEAGRLRSKM-KTMEQIELL-----LOSOLP 178
DB 1481 ONSNKSGETSKDTPIPEDEEERKKVMOQEVLRSLRSIAKELONELRKONOVLODOVK 1540
OY 179 EVEDMIRMGVGSQAVQOLAVYCVSLKREYNLEKARK-----ASG 219
DB 1541 ALQETV-----VSSEASASVHADT-KDLENLKTTEEMLSVTFQVIFNESISDFSTSTA 1594
OY 220 EVADKLKDLFSSRSKQTYISELDQAKLEKSAQKDLQSDAKET-MSLKKKLTMLQETL 278
DB 1595 DFTTFVQKEKREKREILQK-----DVEEOVAVSHQKOLDNIRKLEEMRNKLKLSMLRK- 1647
OY 279 NLPPVASETDRLVLESPAPAYEVNKLRLRPSFRDDIDLNA-----TFPVDT--P 325
DB 1648 NLAVRYALEDOSSKKKSDPAIILSEASKNTDSNKSNSNPAAQVKEKKLIAKTSHTVINS 1707
OY 326 PARPSS 331
DB 1708 PKRSSS 1713

```

```

ID MYHA_HUMAN STANDARD; PRT; 1976 AA.
AC P35580;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-SEP-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, nonmuscle type B (Cellular myosin heavy chain, type B) (Nonmuscle myosin heavy chain-B) (NMHC-B).
GN MYH10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=66025307; PubMed=7499478;
RA Phillips C.L., Yamakawa K., Adelstein R.S.;
RT "Cloning of the cDNA encoding human nonmuscle myosin heavy chain-B and analysis of human tissues with isoform-specific antibodies."
RL J. Muscle Res. Cell Motil. 16:379-389(1995).
RN [2]
RP SEQUENCE OF 63-722 FROM N.A.
RX MEDLINE=91316803; PubMed=1860190;
RA Simons M., Wang M., McBride O.W., Kawamoto S., Yamakawa K., Gula D., Adelstein R.S., Weir L.;
RT "Human nonmuscle myosin heavy chains are encoded by two genes located on different chromosomes."
RL Circ. Res. 69:530-539(1991).
CC -!- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS, CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND CAPING.
CC -!- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M69181; AAA99177.1; -.
DR PIR: B61231; B61231.
DR HSSP: P08799; ILVK.
DR MIM: 160776; -.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KW Myosin; ATP-binding; Calmodulin-binding; Actin-binding; Coiled coil; Alkylation; Multigene family.
FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
FT DOMAIN 786 815 IQ.
FT DOMAIN 845 1976 COILED COIL (POTENTIAL).
FT NP_BIND 178 185 ATP (POTENTIAL).
FT MOD_RES 701 701 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 711 711 ALKYLATION (SH-2) (POTENTIAL).

```

SQ SEQUENCE 1976 AA; 228938 MW; B2B87FF35EA124F CRC64;
 Query Match 6.9%; Score 163.5; DB 1; Length 1976;
 Best Local Similarity 22.78; Pred. No. 0.5;
 Matches 113; Conservative 77; Mismatches 152; Indels 155; Gaps 21;
 QY 35 IQSFETAPSPFCPOCRIOVGKRTIINKLFPLAOEENVLDRP--LKNELDNVRAQLSQ 92
 DB 1387 IESLEEKAKLLKDAE-ALSGRLSEKALAYKLEKTKNRLOQELDDLTVDDHQRVASN 1445
 QY 93 -KDEKRDQVITDPL-----RDPLF---ERNATVVSLOQALGKA----- 128
 DB 1446 LKKKKKFDOLIAEEKSISARYAEERDRAEAREKETKALSLARALEALEAKEEFERQ 1505
 QY 129 -----EMCST-----LKKOMKYLEGOODETK-----OAOEAG-RL 159
 DB 1506 NKQLRADMEDLMSKDVGVKNVHELEKSKRALBOOVEMRTQLELEDELDQATEDAKLRL 1565
 QY 160 RSKKKT-----EQLLELLLOSLPEVEEMIRMGVGSAAVEQLAVYCVSL 204
 DB 1566 EVNNQAKKAQFERDLQTRDEONEEKKRLLIKQVRELEAELED-----ERQORALAVASK 1619
 QY 205 KKEVENKE-----ARKASGEVADKLK-----DLFS-----SR 233
 DB 1620 KKMEDIDKLEAQIEAANKARDEVITKQRLQAQMKDYQRELEBARASRDEIFAQSESE 1679
 QY 234 SKLOTVSELDQAKLELSAQKDSADKELMSLKKLT-----MLOETLNLPPVAS 285
 DB 1680 KKLKSLSELEIILQOEELASSERARRHAEQERDELADETNSAGSKSLDLKRRLEKRIA 1739
 QY 286 ETVDRLVLESPAPVEVNLKLRPSFRDDIDLVDTVPARSSSOHGYYERLCLEKS 345
 DB 1740 QLEELEEE-----QSMELLNDRFR-----KTLQVDTLNA-----ELAERS 1778
 QY 346 HSPIDQVPKKCKGPKRESQSLGSGCAGPEDELYGAF-----PIEVNAILQ--K 397
 DB 1779 AAKSDNARQOLEKRONELKRLKLT-----OLEGAVKSKFKATISALEAKIGOLEE 1828
 QY 398 OPKRPRESSCSKDVVR 414
 DB 1829 QLEQEKERAKANKLVR 1845
 RESULT 35
 ID DESP_HUMAN STANDARD; PRT; 2871 AA.
 AC P15924; Q14189; O75993; Q9UH4;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Desmoplakin (DP) (250/210 kDa paraneoplastic pemphigus antigen).
 GN DSP.
 OS Homo sapiens (human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SPROUNCE FROM N.A. (ISOFORM DPI).
 RC TISSUE-Forceskin;
 RX MEDLINE=92115697; PubMed=1731325;
 RA Virata M.L.A., Wagner R.M., Parry D.A.D., Green K.J.;
 RT "Molecular structure of the human desmoplakin I and II amino
 RT terminus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:544-548(1992).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM DPI).
 RA Phillips S.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SPROUNCE OF 1120-2871 FROM N.A. (ISOFORM DPI).
 RC TISSUE-Forceskin;
 RX MEDLINE=90153880; PubMed=1689290;

RA Green K.J., Parry D.A.D., Steinert P.M., Virata M.L.A., Wagner R.M.,
 RA Angst B.D., Nilles L.A.;
 RT "Structure of the human desmoplakins. Implications for function in
 RT the desmosomal plaque.";
 RL J. Biol. Chem. 265:2603-2612(1990).
 RN [4]
 RP ERRATUM.
 RX MEDLINE=90361712; PubMed=2391353;
 RA Green K.J., Parry D.A.D., Steinert P.M., Virata M.L.A., Wagner R.M.,
 RA Angst B.D., Nilles L.A.;
 RL J. Biol. Chem. 265:11406-11407(1990).
 RN [5]
 RP SEQUENCE OF 2854-2871 FROM N.A.
 RC TISSUE-Skin;
 RX MEDLINE=20062965; PubMed=10594734;
 RA Whitlock N.V., Ashton G.H., Dopling-Hepenstal P.J., Gratian M.J.,
 RA Keane F.M., Eady R.A.J., McGrath J.A.;
 RT "Striate palmoplantar keratoderma resulting from desmoplakin
 RT haploinsufficiency.";
 RL J. Invest. Dermatol. 113:940-946(1999).
 RN [6]
 RP CHARACTERIZATION.
 RX MEDLINE=98012209; PubMed=9348293;
 RA Kowalczyk A.P., Bornslaeger E.A., Borgwardt J.E., Palke H.L.,
 RA Dhaliwal A.S., Corcoran C.M., Denning M.F., Green K.J.;
 RT "The amino-terminal domain of desmoplakin binds to plakoglobin and
 RT clusters desmosomal cadherin-plakoglobin complexes.";
 RL J. Cell Biol. 139:773-784(1997).
 CC -1- FUNCTION: MAJOR HIGH MOLECULAR WEIGHT PROTEIN OF DESMOSOMES.
 CC INVOLVED IN THE ORGANIZATION OF THE DESMOSOMAL CADHERIN-
 CC PLAKGLOBIN COMPLEXES INTO DISCRETE PLASMA MEMBRANE DOMAINS AND IN
 CC THE ANCHORING OF INTERMEDIATE FILAMENTS TO THE DESMOSOMES.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: INNERMOST PORTION OF THE DESMOSOMAL PLAQUE.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: DPI/DPI (SHOWN HERE) AND
 CC DPI/DP2; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: DPI IS APPARENTLY AN OBLIGATE CONSTITUENT OF
 CC ALL DESMOSOMES; DPI RESIDE PREDOMINANTLY IN TISSUES AND CELLS OF
 CC STRATIFIED ORIGIN.
 CC -1- DOMAIN: THE N-TERMINAL REGION IS REQUIRED FOR LOCALIZATION TO THE
 CC DESMOSOMAL PLAQUE AND INTERACTS WITH THE N-TERMINAL REGION OF
 CC PLAKOPHILIN 1. THE C-TERMINAL REGION INTERACTS WITH INTERMEDIATE
 CC FILAMENTS.
 CC -1- DISEASE: DEFECTS IN DSP ARE A CAUSE OF STRIATE PALMOPLANTAR
 CC KERATODERMA II (PKR2, KPS2 OR SPPK2), CHARACTERIZED BY SKIN
 CC THICKENING IN THE PALMS (LINEAR PATTERN) AND THE SOLES (ISLAND-
 CC LIKE PATTERN) AND FLEXOR ASPECT OF THE FINGERS; AND RARELY BY
 CC ABNORMALITIES OF THE NAILS, THE TEETH AND THE HAIR.
 CC -1- SIMILARITY: CONTAINS 17 PLECTRIN REPEATS.
 CC -1- SIMILARITY: CONTAINS 2 SPECTRIN REPEATS.
 CC -1- SIMILARITY: BELONGS TO THE PLAKIN OR CYTOLINKER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M77830; AAA85135.1; -
 CC EMBL: AL031058; CAI19927.1; -
 CC EMBL: J05211; AAA35766.1; -
 CC EMBL: AF139065; AAF19785.1; -
 CC PIR: A35536; A35536.
 CC HSSP: P01100; IPOS.
 CC MIM: 125647; -
 CC InterPro: IPR001101; Plectin_repeat.
 CC InterPro: IPR002017; Spectrin.
 CC Pfam: PF00681; Plectin_repeat; 20.
 CC SMART: SM00250; PLEC; 17.
 CC SMART: SM00150; SPEC; 1.
 KW Repeat; Coiled coil; Phosphorylation; Cytoskeleton;

KW Structural protein; Alternative splicing.
 FT DOMAIN 1 1056
 FT DOMAIN 1057 1945 CENTRAL FIBROUS ROD DOMAIN.
 FT DOMAIN 1946 2871 GLOBULAR 2.
 FT REPEAT 347 447 SPECTRIN 1.
 FT REPEAT 858 952 SPECTRIN 2.
 FT DOMAIN 1018 1945 SPECTRIN 2.
 FT REPEAT 2009 2045 COILED COIL (POTENTIAL).
 FT REPEAT 2046 2083 PLECTIN 1.
 FT REPEAT 2084 2121 PLECTIN 2.
 FT REPEAT 2122 2159 PLECTIN 3.
 FT REPEAT 2163 2197 PLECTIN 4.
 FT REPEAT 2198 2233 PLECTIN 5.
 FT REPEAT 2251 2288 PLECTIN 6.
 FT REPEAT 2289 2326 PLECTIN 7.
 FT REPEAT 2327 2364 PLECTIN 8.
 FT REPEAT 2365 2402 PLECTIN 9.
 FT REPEAT 2406 2440 PLECTIN 10.
 FT REPEAT 2456 2493 PLECTIN 11.
 FT REPEAT 2507 2544 PLECTIN 12.
 FT REPEAT 2610 2647 PLECTIN 13.
 FT REPEAT 2648 2685 PLECTIN 14.
 FT REPEAT 2724 2761 PLECTIN 15.
 FT REPEAT 2762 2799 PLECTIN 16.
 FT DOMAIN 2824 2847 PLECTIN 17.
 FT DOMAIN 1 584 6 X 4 AA TANDEM REPEATS OF G-S-R-[SR].
 MOD_RES 2849 2849 INTERACTS WITH PLAKOPHILIN 1 AND JUNCTION
 VARSPLIC 1195 1794 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
 CONFLICT 905 905 MISSING (IN ISOFORM DPT1).
 CONFLICT 1120 1120 R -> A (IN REF. 2).
 CONFLICT 2687 2688 D -> R (IN REF. 3).
 SEQUENCE 2871 AA; 331771 MW; 5639587CD469087 CMC64;

Query Match 6.9%; Score 163.5; DB 1; Length 2871;
 Best Local Similarity 20.3%; Pred. No. 0.77;
 Matches 96; Conservative 96; Mismatches 156; Indels 125; Gaps 22;

QY 15 DHSRDVAHICGHTFHQCLIOSEFAPSRTCQCRIQVGRKRTIINKLFPDLQAEENVL 74
 DB 1187 EYENELKAVNNHNEESNLRNKRYT-----EINIKRTY----KEISQKED- 1230
 QY 75 DREFKLNELDNVAQLSQDKERKDSQVIID--TLRDTLEBNATVVSLOALGKAEMLC 132
 DB 1231 DSKNLNRQDL-----RLSRENRDLKDELIVRLDSTILOATEQRRAEENALQOKAGSEIMQ 1286
 QY 133 ST--LKKQKTYLEGOODET---KQAGEAGR-LRSKMKTMEDIELLLQ-----S 175
 DB 1287 KKHLELELKQVMQQRSEDAHRKHQSLEFAKTIQDKNKEIERLKAFFEAKRMEYEN 1346
 QY 176 QLEPEV-----EEMT-----RDMGVQSAVEDQAVY---CVSLKKEYENLKAEARASG 219
 DB 1347 ELSVNRNNYDEEITSLKNOFETEINIKTYTHLTLMQKEEDTSGYRAQIDNLTRENSLS 1406
 QY 220 EVADKLKDLFSSRSKQTYSELDQAK-----LELKA-----Q 254
 DB 1407 EETKRLKNTLTQTTENRAREVEDIQOKATGSEVSQKQKQLEVLRYQVTQMRTEESRYK 1466
 QY 255 KDIQASADKETMSLKKLTMLQETLNPPVASETVYDLVLESPAPVEVNLKLRPSEFDDI 314
 DB 1467 QSDDDAKTITQDKNKEIERKQIDL-----KETNDRKCLE-----DENARLQRYQY--- 1512
 QY 315 DLNATFVDTPPARPSSSSOHGY-----YEKLCLEKSHSPIDQVPRKIKKGPKEESOLS 367
 DB 1513 DLOKANSATETINKLKVQDEQLTRLRIDERYVSQETRYKD-QDITR--FQNSLKELOLO 1569
 QY 368 LGGGSCAGEPDEDELVGAFPIFVRNALIGOKQPKRPSSESCSKDVYRTGFDGL 420
 DB 1570 -----KQKVEEEL-----NRKRTASDESGCKRKKLLEEELEGM 1601

RESULT 36

MYH7_HUMAN
 ID MYH7_HUMAN STANDARD: PRT; 1935 AA.
 AC P12883; O14904; O16579;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, cardiac muscle beta isoform (MYHC-beta).
 GN MYH7 OR MYCIB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91065634; PubMed=2249844;
 RA Jaenick T., Diederich K.W., Haas W., Schleich J., Lichter P.,
 RA Florid M., Bach A., Vosberg H.P.;
 RT "The complete sequence of the human beta-myosin heavy chain gene and
 RT a comparative analysis of its product.";
 RL Genomics 8:194-206(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90301496; PubMed=2362820;
 RA Liew C.-C., Sole M.J., Yamauchi-Takihara K., Kellam B.,
 RA Anderson D.H., Lin L., Liew J.;
 RT "Complete sequence and organization of the human cardiac beta-myosin
 RT heavy chain gene.";
 RL Nucleic Acids Res. 18:3647-3651(1990).
 RN [3]
 RP SEQUENCE OF 1-115 FROM N.A.
 RX MEDLINE=89264452; PubMed=2726733;
 RA Yamauchi-Takihara K., Sole M.J., Liew J., Ing D., Liew C.-C.;
 RT "Characterization of human cardiac myosin heavy chain genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3504-3508(1989).
 RN [4]
 RP ERRATUM.
 RA Yamauchi-Takihara K., Sole M.J., Liew J., Ing D., Liew C.-C.;
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7416-7417(1989).
 RN [5]
 RP SEQUENCE OF 1310-1935 FROM N.A.
 RX MEDLINE=86176778; PubMed=2421254;
 RA Saez L., Leinwand L.A.;
 RT "Characterization of diverse forms of myosin heavy chain expressed in
 RT adult human skeletal muscle.";
 RL Nucleic Acids Res. 14:2951-2969(1986).
 RN [6]
 RP REVISIONS.
 RA Leinwand L.A.;
 RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 1410-1935 FROM N.A.
 RX MEDLINE=88299163; PubMed=2969919;
 RA Kurabayashi M., Tsuchimochi H., Komuro I., Takaku F., Yazaki Y.;
 RT "Molecular cloning and characterization of human cardiac alpha- and
 RT beta-form myosin heavy chain complementary DNA clones. Regulation of
 RT expression during development and pressure overload in human
 RT atrium.";
 RL J. Clin. Invest. 82:524-531(1988).
 RN [8]
 RP SEQUENCE OF 785-1935 FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=90235862; PubMed=1691980;
 RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goede H.W.,
 RA Arnold H.H.;
 RT "Identification of three developmentally controlled isoforms of human
 RT myosin heavy chains.";
 RL Eur. J. Biochem. 189:55-65(1990).
 RN [9]
 RP SEQUENCE OF 1393-1935 FROM N.A.
 RX MEDLINE=87192738; PubMed=3032769;
 RA Jandreski M.A., Liew C.-C.;
 RT "Construction of a human ventricular cDNA library and
 RT characterization of a beta myosin heavy chain cDNA clone.";


```

Db 1312 LEDLRKRLSEEVAKNALAHALQASRDHCDLRFQYEETEARKEIQLRVLSKANSEVAQW 1371
Oy 233 RSKLOT-----VSELDQAKLELSAQKDLSQAKKIMSLKKKLTLMDET 277
Db 1372 RTVETDAIORTETELEAKKRL-AQR-LQEAEEAVEAVNAKCSLEKT 1417

RESULT 37
MYSS_CHICK
ID MYSS_CHICK STANDARD; PRT; 1938 AA.
AC P1358; 013228;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Myosin heavy chain, skeletal muscle, adult.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=Pectoralis muscle;
RA Chao T.H., Bandman E., Moore L.;
RT "Cloning, nucleotide sequence and characterization of a full-length
RT cDNA encoding the myosin heavy chain from adult chicken pectoralis
RT major muscle";
RN Submitted (Feb-1997) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE OF 1-205.
RC TISSUE=Pectoralis muscle;
RA Hayashida M., Maeta T., Matsuda G.;
RT "The primary structure of skeletal muscle myosin heavy chain: I.
RT Sequence of the amino-terminal 23 kDa fragment.";
RL J. Biochem. 110:54-59(1991).
RN [3]
RP SEQUENCE OF 206-636.
RC TISSUE=Pectoralis muscle;
RX MEDLINE=92041768; PubMed=1939028;
RA Komine Y., Maeta T., Matsuda G.;
RT "The primary structure of skeletal muscle myosin heavy chain: II.
RT Sequence of the 50 kDa fragment of subfragment-1.";
RL J. Biochem. 110:60-67(1991).
RN [4]
RP SEQUENCE OF 637-837.
RC TISSUE=Pectoralis muscle;
RX MEDLINE=92041769; PubMed=1939029;
RA Maeta T., Miyashita T., Matsuzono K., Tanioka Y., Matsuda G.;
RT "The primary structure of skeletal muscle myosin heavy chain: III.
RT Sequence of the 22 kDa fragment and the alignment of the 23 kDa, 50
RT kDa, and 22 kDa fragments.";
RL J. Biochem. 110:68-74(1991).
RN [5]
RP SEQUENCE OF 838-1938.
RC TISSUE=Pectoralis muscle;
RX MEDLINE=92041770; PubMed=1939030;
RA Maeta T., Yajima E., Nagata S., Miyashita T., Nakayama S., Matsuda G.;
RT "The primary structure of skeletal muscle myosin heavy chain: IV.
RT Sequence of the rod, and the complete 1,938-residue sequence of the
RT heavy chain.";
RL J. Biochem. 110:75-87(1991).
RN [6]
RP PRELIMINARY SEQUENCE OF 1-808.
RX MEDLINE=97092420; PubMed=3467365;
RA Maeta T., Hayashida M., Tanioka Y., Komine Y., Matsuda G.;
RT "The primary structure of the myosin head.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:416-420(1987).
RN [7]
RP SEQUENCE OF 842-1270.
RX MEDLINE=90121764; PubMed=2610940;
RA Watanabe B.;

```

```

RT "Complete amino-acid sequence of subfragment-2 in adult chicken
RT skeletal muscle myosin.";
RL Biol. Chem. Hoppe-Seyler 370:1027-1034(1989).
RN [8]
RP SEQUENCE OF 852-1108.
RX MEDLINE=89374803; PubMed=2775482;
RA Watanabe B.;
RT "Amino-acid sequence of the short subfragment-2 in adult chicken
RT skeletal muscle myosin.";
RL Biol. Chem. Hoppe-Seyler 370:549-558(1989).
RN [9]
RP SEQUENCE OF 1145-1270.
RX MEDLINE=89228549; PubMed=2713098;
RA Watanabe B.;
RT "Amino-acid sequence of the hinge region in chicken myosin
RT subfragment-2.";
RL Biol. Chem. Hoppe-Seyler 370:55-61(1989).
RN [10]
RP SEQUENCE OF 1857-1938 FROM N.A.
RX MEDLINE=87217964; PubMed=3034534;
RA Moriarty D.M., Barringer K.J., Dodgson J.B., Richter H.E.,
RA Young R.B.;
RT "Genomic clones encoding chicken myosin heavy-chain genes.";
RL DNA 6:91-99(1987).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 1-843.
RX MEDLINE=93303624; PubMed=8316857;
RA Rayment I., Rypniewski W.R., Schmidt-Base K., Smith R.,
RA Tomchick D.R., Benning M.M., Winkelman D.A., Wesenberg G.,
RA Holden H.M.;
RT "Three-dimensional structure of myosin subfragment-1: a molecular
RT motor.";
RL Science 261:50-58(1993).
RN [12]
RP FUNCTION: MUSCLE CONTRACTION. MYOSIN IS A PROTEIN THAT BINDS TO
RP F-ACTIN AND HAS ATPASE ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
RN SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
RN HEAVY CHAIN SUBUNITS (MHC), 2 ALKALIC LIGHT CHAIN SUBUNITS (MLC)
RN AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
RN [13]
RP SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
RN DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
RN CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
RN CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
RN [14]
RP PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
RN ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
RN [15]
RP MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
RN MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
RN SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
RN SUBFRAGMENT (S2).
RN [16]
RP SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
RN [17]
RP SIMILARITY: CONTAINS 1 IQ DOMAIN.
RN [18]
RP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement. (See http://www.isb-sib.ch/announce/
RN or send an email to license@isb-sib.ch.)
RN [19]
RP EMBL: U07231; AAB47555.1; -
RN EMBL: M16557; AAA48970.1; -
RN PIR: PX0050; PX0051.
RN PIR: A26821; A26821.
RN PIR: S02082; S02082.
RN PIR: S04501; S04501.
RN PIR: S05515; S05515.
RN PDB: 2MYS; 1JAN-97.
RN InterPro: IPR000048; IO.
RN InterPro: IPR004009; Myosin_N.
RN InterPro: IPR002928; Myosin_Tail.
RN InterPro: IPR001609; myosin_head.
RN Pfam: PF00612; IO; 2.
RN Pfam: PF00063; myosin_head; 1.
DR

```

DR Pfam: PF02736; Myosin_N. 1.
 DR Pfam: PF01576; Myosin_tail. 1.
 DR Prodom: PD000355; myosin_head. 1.
 DR SMART: SM00015; IQ. 1.
 DR SMART: SM00242; MYSC. 1.
 DR PROSITE: PS50096; IQ. 1.
 DR Myosin: Muscic protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Phosphorylation; Acetylation;
 KM Calmodulin-binding; Multigene family; 3D-structure.
 FT INT_MET 0
 FT DOMAIN 1 782 MYOSIN HEAD-LIKE.
 FT DOMAIN 783 812 IQ.
 FT DOMAIN 838 840 HINGE.
 FT DOMAIN 841 1938 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP (POTENTIAL).
 FT DOMAIN 657 679 ACTIN-BINDING.
 FT DOMAIN 759 773 ACTIN-BINDING.
 FT MOD_RES 1 1 ACETYLATION.
 FT MOD_RES 35 35 METHYLATION (MONO-).
 FT MOD_RES 130 130 METHYLATION (TRI-).
 FT MOD_RES 551 551 METHYLATION (TRI-).
 FT MOD_RES 755 755 METHYLATION (MONO-).
 FT MOD_RES 697 697 ALKYLATION (SH-1).
 FT MOD_RES 707 707 ALKYLATION (SH-2).
 FT CONFLICT 907 907 C -> Q (IN REF. 7 AND 8).
 FT CONFLICT 980 980 L -> F (IN REF. 1).
 FT CONFLICT 1343 1343 E -> D (IN REF. 5).
 FT CONFLICT 1545 1545 S -> A (IN REF. 5).
 FT CONFLICT 1796 1797 H V -> Q (IN REF. 5).
 FT CONFLICT 1830 1830 S -> A (IN REF. 5).
 FT CONFLICT 1863 1863 I -> V (IN REF. 10).
 FT CONFLICT 1929 1931 IHG -> FH (IN REF. 10).
 SQ SEQUENCE 1938 AA; 223013 MW; EDD01CEA2681E10F CRC64;

Query Match 6.88; Score 163; DB 1; Length 1938;
 Best Local Similarly 22.34; Pred. No. 0.52;
 Matches 65; Conservative 58; Mismatches 107; Indels 62; Gaps 9;

QY 31 LQCIQSFETAPSRCTPCQCRIOVGKRTIINKLFFDLAONEENVLDREF----- 78
 DB 1145 LEEISERLEEGCATTAAGIENKKEAEFQMRDL--EATLQHEHTAALAKKKHADQ 1201
 QY 79 ---LKNELDNVRAQLSQKDEKRDQYIDLPRLTEERNATVVSLOQALGKAMLCSTL 135
 DB 1202 TAEIGEDIDNQRVKQKLEKSKELMEIDLASME-----SVKAKANLEKMCRTL 1254
 QY 136 KKQKKYLEQOODE-----TKQAEAGRL-----RSKMKIMEQI 169
 DB 1255 EDOLSEIKTEEQNQRMINDLTQRLQETGEGYSQAEEKDALLISQSKGCGFTQOI 1314
 QY 170 ELLIQLSLPEEEMIRMGVGSVEQLAVYCVSLKKEYEMLKARKASGVADKLKLDL 229
 DB 1315 EELKR---HLEERIKKMLAHLAQSRHDCELRLREYEEQDAKGELORALSKANSEV 1370
 QY 230 FSSRSKIQT---VYSELDOAKLELSAQKDLQSDAKREISLKKLMLQET 277
 DB 1371 AQMTKRYETDAIQTEELBEAKKLL--AQK-LQDAEHEVAENNAKCSLEKT 1419

RESULT 38
 KFAA_HUMAN STANDARD: PRT; 1232 AA.
 ID KFAA_HUMAN 095239; 09NNT6; 09NY24;
 AC 15-JUL-1999 (Rel. 38, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Chromosome-associated kinesin KIF4A (Chromokinesin).
 GN KIF4A OR KIF4.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;

RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE-Lymphocytes;
 RA Villard L.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RC SEQUENCE FROM N.A.; AND TISSUE SPECIFICITY.
 RP MEDLINE=20435301; Pubmed=10978527;
 RA Oh S.J., Hahn H., Torrey T.A., Shin H., Choi W., Lee Y.M.,
 RA Morse H.C., Iii, Kim W.;
 RT "Identification of the human homologue of mouse Kif4, a kinesin
 superfamily motor protein.";
 RL Biochim. Biophys. Acta 1493:219-224(2000).
 RN [3]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Retinoblastoma;
 RA Kentsch A., Neumann T., Rommerskirch W.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RC SEQUENCE OF 128-1232 FROM N.A.
 RP TISSUE=Retinoblastoma;
 RA MEDLINE=97311419; Pubmed=9168136;
 RA van R.-T., Wang S.-Z.;
 RT "Increased chromokinesin immunoreactivity in retinoblastoma cells.";
 RL Gene 189:263-267(1997).
 CC -1- FUNCTION: REQUIRED FOR MITOTIC CHROMOSOMAL POSITIONING AND BIPOlar
 SPINDLE STABILIZATION (BY SIMILARITY)
 CC -1- SUBCELLULAR LOCATION: NUCLEAR, ASSOCIATED WITH MITOTIC
 CHROMOSOMES (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEMATOPOIETIC TISSUES,
 FETAL LIVER, SPLEEN, THYMUS AND ADULT THYMUS AND BONE MARROW.
 CC LOWER LEVELS ARE FOUND IN HEART, TESTIS, KIDNEY, COLON AND LUNG.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
 CC CHROMOKINESIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF179308; AAD51855.1; -
 CC EMBL: AF071592; AAD05492.2; -
 CC EMBL: AJ271784; CAB75427.1; -
 CC EMBL: AF277375; AAF86334.1; -
 CC HSSP: P17119; 3KAR.
 DR InterPro: IPR001752; kinesin.
 DR PRINTS: PR00380; KINESINHEAVY.
 DR SMART: SM00129; KISC. 1.
 DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
 KW Motor protein; Microtubules; ATP-binding; DNA-binding;
 KM Nuclear protein; Coiled coil.
 FT DOMAIN 1 349 KINESIN-MOTOR.
 FT DOMAIN 350 999 COILED COIL (BY SIMILARITY).
 FT DOMAIN 1000 1232 GLOBULAR.
 FT NP_BIND 88 95 ATP (POTENTIAL).
 FT CONFLICT 223 223 R -> G (IN REF. 2).
 FT CONFLICT 231 231 S -> T (IN REF. 4).
 FT CONFLICT 286 286 V -> A (IN REF. 2).
 FT CONFLICT 422 422 L -> W (IN REF. 2).
 FT CONFLICT 564 564 L -> H (IN REF. 4).
 FT CONFLICT 564 564 L -> P (IN REF. 2).
 FT CONFLICT 600 600 R -> E (IN REF. 3).
 FT CONFLICT 668 668 R -> K (IN REF. 3 AND 4).
 FT CONFLICT 928 928 Q -> P (IN REF. 1).
 FT CONFLICT 958 958 QL -> RL (IN REF. 3).
 FT CONFLICT 960 960 L -> Q (IN REF. 1).
 FT CONFLICT 996 997 LL -> S (IN REF. 4).
 FT CONFLICT 1003 1014 OKHLPRDTLLSP -> RLPLRIPFYQL (IN REF. 4).

FT CONFLICT 1022 1022 P -> Q (IN REF. 2).
 FT CONFLICT 1077 1077 K -> N (IN REF. 2).
 FT CONFLICT 1138 1138 G -> S (IN REF. 2).
 SQ SEQUENCE 1232 AA; 139908 MW; FF74052A17A6E8F7 CRC64;

Query Match 6.8%; Score 162.5; DB 1; Length 1232;
 Best Local Similarity 21.4%; Pred. No. 0.33;
 Matches 104; Conservative 67; Mismatches 143; Indels 173; Gaps 21;

QY 65 DLQAEENVDREFLNK-----ELDNVRAQLSQDKKRRDS-QVITDPLRDTLEE 113
 DB 777 DVQLKEK-----KESGPNPPKLRRTFSLTEVREGQVS-----ESEDSTTKQIESLETMEF 829
 QY 114 RNATFVSLQALKAEMLCSTLKQKKMYLEQQQDEFQQAQEF-AGRLRSKMKTKMEQELL 172
 DB 830 RSNQIADLQOKLDAE-----SEDRPKQRMENATITLCAKALKYLIGEL 874
 QY 173 LQSQLEPEVEEMIRDMGQSGAVEQLAVYCVSLKKEYENLKEARKASGEVADKLKRLDF-- 230
 DB 875 VSKRI-QVSKL-----ESLQSKTSKADMOK---MLPEERNHFAFETELQAEIVRM 923
 QY 231 --SSRSKLQTVSELDQAKLELSAQKDLQSAKDEMS----- 266
 DB 924 EQHQEKEVLYLSLOLOSOQMAEKQLEESVSEKEQQLLSTLKCQDELEKKREVCQEQOQL 983
 QY 267 -----LKKRTLMQETLNLPRVASEVVDRLVLESPAPVEVNLKLRPRFRDDIDLNATF 320
 DB 984 LRENEITKQTLTQVASRQKHLPKDT-----LISP-----DSSF 1018
 QY 321 DVDTPPARPSOCHGYE-----KLCLKSHSPIOD-----VPKIKCG 359
 DB 1019 EYVPRKPRPRVKEKLEQGMEDIEDLKCEHSHVNEHEDDGDDEDEMKRTKLKV 1078
 QY 360 PKRESQSLQSGQSCAGPEDELVGAPFIEVFNAILGQKQKPRPSSESSCSKDVV----- 413
 DB 1079 SRKNIQ-----GCSCKG-----WCGNKQCGCRKQKDCGDCDDCDPTKC 1117
 QY 414 ---RTGFDGLG-----GRFKFIPTDPLV---MIRPLPVKPKR-----VKORV 450
 DB 1118 RNRQCGKDSIGYVERTDSSGSKFLEDPTEVTEGSLFFNPVCATPNSKILKMCQVBOVL 1177
 QY 451 RVKTVPS 457
 DB 1178 SKTKPPA 1184

RESULT 39
 MYS_AEOIR
 ID MYS_AEOIR STANDARD; PRT; 1938 AA.
 AC P24733;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, striated muscle.
 OS Aequipecten irradians (Bay scallop).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoida;
 OC Pectinoidae; Pectinidae; Argopecten.
 NC NCB1_taxid=31199;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Adductor muscle;
 RX MEDLINE=92011595; PubMed=1917970;
 RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
 RT "Complete primary structure of a scallop striated muscle myosin heavy
 chain. Sequence comparison with other heavy chains reveals regions
 that might be critical for regulation.";
 RL J. Biol. Chem. 266:18469-18476(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Adductor muscle;
 RX MEDLINE=91088319; PubMed=2263488;
 RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;

RT "Nucleotide sequence of full length cDNA for a scallop striated
 RT muscle myosin heavy chain.";
 RL Nucleic Acids Res. 18:7158-7158(1990).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.
 RX MEDLINE=94173332; PubMed=8127365;
 RA Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalabokis V.N.,
 RT Szent-Gyorgyi A.G., Cohen C.;
 RT "Structure of the regulatory domain of scallop myosin at 2.8-A
 RT resolution.";
 RL Nature 368:306-312(1994).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.
 RX MEDLINE=96419133; PubMed=8805510;
 RA Houdusse A., Cohen C.;
 RT "Structure of the regulatory domain of scallop myosin at 2-A
 RT resolution. Implications for regulation.";
 RL Structure 4:21-32(1996).
 CC CC
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
 CC CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC CC
 CC EMBL; X55714; CAA39247.1; -
 CC PIR; S13557; S13557.
 CC PIR; A40997; A40997.
 CC PDB; 1SCM; 30-APR-94.
 CC PDB; 1WDC; 11-JUL-96.
 CC InterPro: IPR000048; IQ.
 CC InterPro: IPR004009; Myosin_N.
 CC InterPro: IPR002928; Myosin_tail.
 CC InterPro: IPR002017; Spectrin.
 CC InterPro: IPR001609; myosin_head.
 CC Pfam: PF00612; IQ; 2.
 CC Pfam: PF00063; myosin_head; 1.
 CC Pfam: PF02736; Myosin_N; 1.
 CC Pfam: PF01576; Myosin_tail; 1.
 CC PRINTS: PR00193; MYOSINHEAVY.
 CC ProDom: PD000355; myosin_head; 1.
 CC SMART; SM00015; IQ; 1.
 CC SMART; SM00242; MYSC; 1.
 CC POSITIVE; PSS0096; IQ; 1.
 CC DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 CC KW ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.
 CC FT DOMAIN 1 777
 CC FT DOMAIN 778 805
 CC FT DOMAIN 836 1938
 CC FT DOMAIN 836 1938
 CC FT NP_BIND 176 183
 CC FT MOD_RES 693 693
 CC FT MOD_RES 703 703
 CC FT HELIX 778 821
 CC FT TURN 822 823
 CC FT HELIX 825 833
 CC FT TURN 834 835
 CC FT TURN 834 835
 SQ SEQUENCE 1938 AA; 222821 MW; ASCCEA127D1A4896 CRC64;

Query Match 6.8%; Score 162.5; DB 1; Length 1938;
 Best Local Similarity 25.1%; Pred. No. 0.55;

Matches 63; Conservative 55; Mismatches 84; Indels 49; Gaps 9;

OY 51 IOVGKRTIINKLFDPDIAOEENVDREFLNEL-DNVR---AOLSQKREKRSQVIIDT 106
 Db 1045 VEKAKR---KVEDDLKSTOEVEDLEKYKRELEENVRKREKREISLSLNKLEDEONLSQ 1100
 OY 107 LRDPLEERNATVVSILQALGRAEMLCSTLKROMKXYLEQOODETKOAEAGRLRSKMKTM 166
 Db 1101 LQRIKELQARIELELELELEARNARAKVEKQRAELNRELEBELGRLEADG-----ATS 1155
 OY 167 EQIELLLOSLPEVEEMIRDMGVGOSAVEQLAVYCVSLKKEKENLKEKRSKSGEYADKLR 226
 Db 1156 AQELNKKRE-AELTKIRDL---EEASLOHQAQISALRKKHOD-----ANEMADQV- 1204
 OY 227 KDLFSRSKLOTVYSELQOAKLELKSQKDLQSA-----DKREIMSLK 268
 Db 1205 -----DOLQKVSKEKDKKDLKREMDDELSQHTHMKKCKGCEKWKQEPESQMDLN 1257
 OY 269 KKLTMLOETLN 279
 Db 1258 ARLEDSQRSIN 1268

RESULT 40
 ID AKAP9_HUMAN STANDARD; PRT; 3911 AA.
 AC 099996; 09U004; 09U0H3; 09Y6Y2; 014869; 043555; 094895; 09Y6B8;
 DT 16-Oct-2001 (Rel. 40, Created)
 DT 16-Oct-2001 (Rel. 40, Last sequence update)
 DE A-kinase anchor protein 9 (Protein kinase A anchoring protein 9)
 DE (PKA9) (A-kinase anchor protein 450 kDa) (AKAP 450) (A-kinase anchor
 DE protein 350 kDa) (AKAP 350) (hAKAP 350) (AKAP 120 like protein)
 DE (Hyperion protein) (Yotiao protein) (Centrosome- and golgi-localized
 DE PKA-associated protein) (CG-NAP).
 GN AKAP9 OR AKAP450 OR AKAP350 OR KIAA0803.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eularchia; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 4).
 RC TISSUE-Brain;
 RX MEDLINE=98151389; PubMed=9482789;
 RA Lin J.W., Wyszynski M., Madhavan R., Sealock R., Kim J.U., Sheng M.;
 RT "Yotiao, a novel protein of neuromuscular junction and brain that
 RT interacts with specific splice variants of NMDA receptor subunit
 RT NR1.",
 RL J. Neurosci. 18:2017-2027(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND VARIANT GIN-1347 INS.
 RX MEDLINE=99219864; PubMed=10202149;
 RA Wilczek O., Skalsheg B.S., Kerker G., Bornens M., Tasken K.,
 RA Jansen T., Oerstavik S.,
 RT "Cloning and characterization of a cDNA encoding an A-kinase anchoring
 RT protein located in the centrosome, AKAP450.",
 RL EMBO J. 18:1858-1868(1999).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC TISSUE-Brain;
 RX MEDLINE=99287934; PubMed=10358086;
 RA Takahashi M., Shibata H., Shimakawa M., Miyamoto M., Mukai Y.,
 RT "Characterization of a novel giant scaffolding protein, CG-NAP, that
 RT anchors multiple signaling enzymes to centrosome and the golgi
 RT apparatus.",
 RL J. Biol. Chem. 274:17267-17274(1999).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Kemmer W.A., Deiss S., Schwarz U.;
 RT "Cloning of Hyperion".
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 323-3911 FROM N.A. (ISOFORM 2).

RC TISSUE-Gastric parietal cell;
 RX MEDLINE=99115654; PubMed=9915845;
 RA Schmidt P.H., Dransfield D.T., Claudio J.O., Hawley R.G.,
 RA Trotter K.W., Milgram S.L., Goldring J.R.;
 RT "AKAP350, a multiply spliced protein kinase A-anchoring protein
 RT associated with centrosomes",
 RL J. Biol. Chem. 274:3055-3066(1999).
 RN [6]
 RP SEQUENCE OF 1802-3876 FROM N.A. (ISOFORM 5).
 RC TISSUE-Lymphoblast;
 RA Hinds K., Sutterer C., Becker M., Hawkins M.;
 RT Submitted (Jan-1998) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 2157-3911 FROM N.A. (ISOFORM 6).
 RC TISSUE-Lung;
 RA Milgram S.L., Goldring J.R., Schmidt P.H.;
 RT "AKAP350: A multiply spliced family of proteins with centrosomal
 RT association.",
 RL Submitted (Sep-1998) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE OF 2212-3911 FROM N.A. (ISOFORM 2/3).
 RC TISSUE-Brain;
 RX MEDLINE=99087487; PubMed=9872452;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
 RA Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XI.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro",
 RL DNA Res. 5:277-286(1998).
 RN [9]
 RP SEQUENCE OF 17-1800 FROM N.A.
 RA Wu X., Graves T., Bradshaw H.;
 RT Submitted (Sep-1998) to the EMBL/GenBank/DBJ databases.
 RL [10]
 RP FUNCTION: BINDS TO TYPE II REGULATORY SUBUNIT OF PROTEIN KINASE
 CC A. SCAFFOLDING PROTEIN THAT ASSEMBLES SEVERAL PROTEIN KINASES AND
 CC PHOSPHATASES ON CENTROSOME AND GOLGI APPARATUS WHERE PHYSIOLOGICAL
 CC EVENTS CAN BE REGULATED BY PHOSPHORYLATION STATE OF PROTEIN
 CC SUBSTRATES. ISOFORM 4/YOTIAO IS ASSOCIATED WITH THE N-METHYL-D-
 CC ASPARTATE RECEPTOR AND IS SPECIFICALLY FOUND IN THE NEUROMUSCULAR
 CC JUNCTION (NMJ) AS WELL AS IN NEURONAL SYNAPSES EXPLAINING THAT ITS
 CC ROLE MAY BE TO ORGANIZE POSTSYNAPTIC SPECIALIZATIONS.
 CC [11]
 CC SUBUNIT: INTERACTS WITH THE REGULATORY REGION OF PROTEIN KINASE N
 CC (PKN), PROTEIN PHOSPHATASE 2A (PP2A), PROTEIN PHOSPHATASE 1 (PPI)
 CC AND THE IMATURE NON-PHOSPHORYLATED FORM OF PKC EPSILON.
 CC [12]
 CC SUBCELLULAR LOCATION: CENTROSOMAL IN MANY CELL TYPES AND
 CC CYTOPLASMIC IN PARIETAL CELLS.
 CC [13]
 CC ALTERNATIVE PRODUCTS: 6 ISOFORMS: 1 (SHOWN HERE), 2, 3/CG-
 CC NAP, 4/YOTIAO, 5 AND 6/AKAP350; ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC [14]
 CC TISSUE SPECIFICITY: WIDELY EXPRESSED. ISOFORM 4/YOTIAO IS HIGHLY
 CC EXPRESSED IN SKELETAL MUSCLE AND IN PANCREAS.
 CC [15]
 CC DOMAIN: RII BINDING SITE, PREDICTED TO FORM AN AMPHIPATHIC HELIX,
 CC COULD PARTICIPATE IN PROTEIN-PROTEIN INTERACTIONS WITH A
 CC COMPLEMENTARY SURFACE ON THE R-SUBUNIT DIMER.
 CC [16]
 CC CAUTION: REF. 6 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO TWO
 CC FRAMESHIFTS IN POSITIONS 3782 AND 3811.
 CC [17]
 CC CAUTION: REF. 9 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FOUR
 CC FRAMESHIFTS IN POSITIONS 29, 1653, 1699 AND 1735.
 CC [18]
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC [19]
 CC EMBL: AJ131693; CAB40713.1;
 CC EMBL: AB019691; BAA78718.1;
 CC EMBL: A010770; CA009361.1;
 CC EMBL: AF026245; AAB86384.1;
 CC EMBL: AF083037; AAD22767.1;
 CC EMBL: AC004013; AAB96867.1; ALT_FRAME.

DR EMBL: AF091711; AAD39719.1; -
DR EMBL: AB018346; BAA34523.1; -
DR EMBL: AC000066; AAC60380.1; ALT_FRAME.
DR MTM: 604001; -
KW Coiled coil; Alternative splicing; Polymorphism.
FT DOMAIN 2554 2567 PKA-RIT SUBUNIT BINDING DOMAIN.
FT DOMAIN 164 914 COILED COIL (POTENTIAL).
FT DOMAIN 944 1022 COILED COIL (POTENTIAL).
FT DOMAIN 1100 1185 COILED COIL (POTENTIAL).
FT DOMAIN 1253 1280 COILED COIL (POTENTIAL).
FT DOMAIN 1336 1392 COILED COIL (POTENTIAL).
FT DOMAIN 1434 1459 COILED COIL (POTENTIAL).
FT DOMAIN 1585 1659 COILED COIL (POTENTIAL).
FT DOMAIN 1857 2455 COILED COIL (POTENTIAL).
FT DOMAIN 2544 2561 COILED COIL (POTENTIAL).
FT DOMAIN 2603 2776 COILED COIL (POTENTIAL).
FT DOMAIN 3065 3092 COILED COIL (POTENTIAL).
FT DOMAIN 3124 3470 COILED COIL (POTENTIAL).
FT DOMAIN 3587 3689 COILED COIL (POTENTIAL).
FT DOMAIN 3726 3730 COILED COIL (POTENTIAL).
FT DOMAIN 203 292 POLY-LEU.
FT DOMAIN 321 1010 GLN-RICH.
FT DOMAIN 1846 2772 GLU-RICH.
FT VARSPPLIC 17 28 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPPLIC 1637 1642 OLQEEI -> LATRD (IN ISOFORM 4).
FT VARSPPLIC 1643 3911 MISSING (IN ISOFORM 4).
FT VARSPPLIC 2175 2182 MISSING (IN ISOFORM 3).
FT VARSPPLIC 2175 2183 SADTFORVE -> Q (IN ISOFORM 6).
FT VARSPPLIC 2895 2907 VEGFYNNCESTIC -> GSSIPPLAHSADAYOTREISS
(IN ISOFORM 2, ISOFORM 3 AND ISOFORM 6).
FT VARSPPLIC 2895 2948 MISSING (IN ISOFORM 5).
FT VARSPPLIC 2901 3911 STIOFHAGMR -> ALSLTTSWQHHSAPRTAPLFEILLSH
SIG (IN ISOFORM 6).
FT VARIANT 1347 1347 K -> KO.
FT CONFLICT 76 76 /FTID=VAR_010926.
FT CONFLICT 475 475 E -> Q (IN REF. 3).
FT CONFLICT 554 554 M -> I (IN REF. 3).
FT CONFLICT 638 638 E -> G (IN REF. 3).
FT CONFLICT 663 663 R -> S (IN REF. 3).
FT CONFLICT 913 913 N -> S (IN REF. 3).
FT CONFLICT 956 956 H -> N (IN REF. 3).
FT CONFLICT 980 982 K -> N (IN REF. 3).
FT CONFLICT 997 997 OKH -> PKP (IN REF. 1 AND 2).
FT CONFLICT 1001 1001 Q -> P (IN REF. 1 AND 2).
FT CONFLICT 1020 1020 N -> D (IN REF. 3).
FT CONFLICT 1028 1028 V -> E (IN REF. 3).
FT CONFLICT 1626 1626 R -> P (IN REF. 3).
FT CONFLICT 1703 1703 N -> T (IN REF. 3).
FT CONFLICT 1707 1707 V -> G (IN REF. 3).
FT CONFLICT 1802 1803 MISSING (IN REF. 5).
FT CONFLICT 1843 1843 A -> P (IN REF. 3).
FT CONFLICT 1956 1956 I -> V (IN REF. 3).
FT CONFLICT 2027 2027 V -> D (IN REF. 5).
FT CONFLICT 2157 2158 EI -> HE (IN REF. 7).
FT CONFLICT 2169 2169 E -> V (IN REF. 3).
FT CONFLICT 2514 2514 L -> R (IN REF. 3).
FT CONFLICT 2851 2851 I -> N (IN REF. 8).
FT CONFLICT 2957 2957 E -> D (IN REF. 3).
FT CONFLICT 2983 2983 P -> S (IN REF. 3, 7 AND 8).
FT CONFLICT 3087 3087 Q -> H (IN REF. 3).
FT CONFLICT 3218 3218 Q -> H (IN REF. 3).
FT CONFLICT 3307 3309 ESE -> OSO (IN REF. 3).
FT CONFLICT 3751 3751 P -> A (IN REF. 3).
FT CONFLICT 3833 3833 T -> S (IN REF. 3).
SQ SEQUENCE 3911 AA; 453664 MM; 3FB1CB1C819B47AA CRC64;

Db 1940 ADEKTLFERQIQ-EKTDIIDRLQELLCSNRLQELAEQOQIQEERELLRSQKEMKME 1998
Qy 83 LDVVRQALSQKD-----KEKDSOVIITDRLDTL-----BERNA 116
Db 1999 AGPEEQDLQETKLMEKELEVOCAEKVRDLDQKVALEIDVEBOVSRTIELDEBKVN 2058
Qy 117 TVVSLQOALGKAMLCSTLKKOMKYLEQO-----QDETQAOEAGRLRSKMTQETLEL 172
Db 2059 ELMDLRO---QNALQKLEKMKRFLDEQALIDREHERDVQOELQKLEQALKVPPRPPI 2115
Qy 173 LOSQLPEVEEMIRDMGVGOSAVBOLAVYCVSLKKE-YENLKEARKASGEVADRLRKDLS 231
Db 2116 SEHQTRVEQDLAHMLKKTCKSCCELLSKRQQLQDQGERNELEKLEFVRRE-LEQALLV 2174
Qy 232 SRSKLQTVSELD---QAKLELK-----SAQKD-LQSADEKIMSLKKITLMLOETLNP 281
Db 2175 SADTFORVEDRKRHFGAWEAKPELSLEVOLOAERDAIDRKEKEITNLEEQLEOFREBLE-- 2232
Qy 282 PVASETVDRVLV-----ESPAVEVNLKLRPSFPDDIDLNATPDVDPAPRPSSSH 334
Db 2233 -NKNEEVQQLHMOLEIQKKESTTRIQ-ELQENKLRKQME-KLGALAKESDAMSTODH 2289
Qy 335 GYEEKLC-LEKSHSPIQDVPKIKCKGPKRESQLSLGOSCAGEPDELYGAFPIFVRNA 392
Db 2290 VLGKFKQIIOEKKEVEIDQLEQVTK---LQQLKITTDKVIIEKKNELIRDETQIECL 2346
Qy 393 ILGOKPKPRPSE 405
Db 2347 MSDQECVKNRNE 2359

Search completed: September 4, 2002, 16:17:19
Job time: 478 sec

Query Match 6.8%; Score 162.5; DB 1; Length 3911;
Best Local Similarity 22.2%; Pred. No. 1.2; Mismatches 81; Gaps 18;
Matches 96; Conservative 80; Indels 176; Indels 81; Gaps 18;
Qy 41 APSRTCPCRCRIQGVKRTINKLFPDL-----AQEEENVLDLR--EFLKNE 82
